

GenCore version 5.1.4-p5-4578  
 Copyright (c) 1993 - 2003 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: May 27, 2003, 16:04:16 ; Search time 17 Seconds

(without alignment)  
 655,958 million cell updates/sec

Title: US-09-944-457-2

Perfect score:

1992

Sequence:

1 MKEYVILLFLAICSAKPFES.....PATFRCVLSRMSVQLGNFGM

379

Scoring table: BloSUM62

Gapop 10.0 , Gapext 0.5

Scored: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 500 summaries

Database : Issued\_Patents\_AA:\*

1: /cn2\_6/p2odata/1/iaa/5A\_COMB.pep:\*

2: /cn2\_6/p2odata/1/iaa/5B\_COMB.pep:\*

3: /cn2\_6/p2odata/1/iaa/6A\_COMB.pep:\*

4: /cn2\_6/p2odata/1/iaa/6B\_COMB.pep:\*

5: /cn2\_6/p2odata/1/iaa/PCUTS\_COMB.pep:\*

6: /cn2\_6/p2odata/1/iaa/backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No. Score Query Match Length DB ID Description

1 1799 90.3 373 4 US-09-724-864-43 Sequence 43, Appli

2 1021 51.3 368 1 US-08-303-238-3 Sequence 3, Appli

3 1021 51.3 368 1 US-08-458-834-3 Sequence 3, Appli

4 1020 51.2 368 6 5340934-2- Sequence 2, Appli

5 963.5 48.4 359 1 US-08-303-238-4 Sequence 4, Appli

6 963.5 48.4 359 4 US-08-452-834-4 Sequence 4, Appli

7 949 47.6 342 1 US-08-272-919-2 Sequence 2, Appli

8 949 47.6 342 1 US-08-619-916-2 Sequence 2, Appli

9 949 47.6 342 1 PCT-US95-08542-2 Sequence 27, Appli

10 947 47.5 333 1 US-08-442-063A-27 Sequence 2, Appli

11 907 45.5 353 6 5340934-4 Sequence 2, Appli

12 847 42.5 307 1 US-08-442-063A-48 Sequence 48, Appli

13 776 39.0 282 1 US-08-442-063A-45 Sequence 45, Appli

14 616 30.9 236 1 US-08-442-063A-42 Sequence 42, Appli

15 493 24.7 188 1 US-08-442-063A-39 Sequence 39, Appli

16 384.5 19.3 141 1 US-08-442-063A-36 Sequence 36, Appli

17 342.5 17.2 375 1 US-08-303-238-2 Sequence 2, Appli

18 34.5 17.2 375 4 US-08-458-834-2 Sequence 2, Appli

19 336 16.9 1525 3 US-09-191-647-2 Sequence 2, Appli

20 336 16.9 1525 4 US-09-540-245A-2 Sequence 2, Appli

21 336 16.9 1525 4 US-09-540-153-2 Sequence 2, Appli

22 334.5 16.8 649 4 US-09-188-930-305 Sequence 305, Appli

23 318 16.0 376 1 US-08-303-238-1 Sequence 1, Appli

24 318 16.0 376 4 US-08-458-834-1 Sequence 1, Appli

25 316 15.9 1480 3 US-09-191-647-2 Sequence 7, Appli

26 316 15.9 1480 4 US-09-540-245A-7 Sequence 7, Appli

27 316 15.9 1480 4 US-09-540-153-8 Sequence 7, Appli

28 316 15.9 1480 5 US-09-182-024A-5 Sequence 2, Appli

29 313 15.7 1480 4 US-09-182-024A-5 Sequence 2, Appli

30 309 15.5 1523 4 US-09-182-024A-2 Sequence 5, Appli

31 309 15.5 1091 3 US-08-986-885-5 Sequence 5, Appli

32 272 14.9 603 1 US-08-190-802A-50 Sequence 50, Appli

33 272 13.7 603 4 US-08-477-889-50 Sequence 50, Appli

34 272 13.7 603 4 US-08-473-889-50 Sequence 50, Appli

35 272 13.7 603 4 US-08-487-202A-50 Sequence 50, Appli

36 262 13.2 708 4 US-09-131-048-2 Sequence 2, Appli

37 259.5 11.0 1101 3 US-08-086-885-2 Sequence 2, Appli

38 256.5 12.9 605 4 US-08-066-886-33 Sequence 33, Appli

39 254 12.8 353 3 US-08-986-885-6 Sequence 6, Appli

40 246 12.3 605 4 US-08-190-002A-49 Sequence 49, Appli

41 246 12.3 605 4 US-08-473-089-49 Sequence 49, Appli

42 246 12.3 440 4 US-09-410-372-3 Sequence 49, Appli

43 246 12.3 605 4 US-08-487-072A-49 Sequence 4, Appli

44 241 12.1 605 4 US-09-063-950-5 Sequence 5, Appli

45 241 12.1 673 4 US-09-063-950-2 Sequence 2, Appli

46 239 12.0 180 3 US-08-986-885-8 Sequence 8, Appli

47 236 11.8 440 3 US-08-985-335-3 Sequence 3, Appli

48 236 11.8 440 4 US-09-410-372-3 Sequence 3, Appli

49 222 11.7 298 4 PCT-US91-03055-4 Sequence 4, Appli

50 212.5 10.7 600 4 US-09-232-160-17 Sequence 17, Appli

51 206 10.3 222 5 PCT-US91-09055-3 Sequence 3, Appli

52 206 10.3 231 3 US-08-986-485-7 Sequence 7, Appli

53 205 10.3 313 3 US-08-985-335-8 Sequence 8, Appli

54 205 10.3 313 4 US-09-410-372-8 Sequence 8, Appli

55 199.5 10.0 560 3 US-08-592-500-2 Sequence 2, Appli

56 199.5 10.0 560 3 US-08-195-006-2 Sequence 2, Appli

57 199.5 10.0 560 4 US-09-063-950-4 Sequence 2, Appli

58 199.5 10.0 560 4 US-09-063-950-2 Sequence 2, Appli

59 195 9.8 610 1 US-08-702-717B-6 Sequence 6, Appli

60 195 9.8 610 1 US-08-119-262B-6 Sequence 6, Appli

61 195 9.8 610 1 US-08-135-029A-11 Sequence 11, Appli

62 195 9.8 610 1 US-08-234-065A-11 Sequence 11, Appli

63 194 9.7 320 1 US-07-613-083B-1 Sequence 1, Appli

64 184.5 9.3 196 5 PCT-US91-03055-6 Sequence 6, Appli

65 181 9.1 746 5 PCT-US95-10509-2 Sequence 5, Appli

66 177 8.9 196 5 PCT-US91-09055-5 Sequence 4, Appli

67 175.5 8.8 661 1 US-08-314-014-4 Sequence 4, Appli

68 175 8.8 661 2 US-08-833-823-4 Sequence 4, Appli

69 175 8.7 551 1 US-08-442-063A-57 Sequence 57, Appli

70 173 8.7 551 1 US-08-153-085-3 Sequence 3, Appli

71 172 8.6 1112 4 US-09-153-085-2 Sequence 3, Appli

72 170.5 8.6 1112 4 US-09-188-090-124 Sequence 124, Appli

73 170.5 8.6 806 4 US-08-945-083-2 Sequence 2, Appli

74 169 8.5 1016 4 US-09-180-139-8 Sequence 8, Appli

75 168 8.4 968 4 US-09-180-1439-3 Sequence 6, Appli

76 168 8.4 968 4 US-09-180-139-4 Sequence 4, Appli

77 166.5 8.4 134 3 US-09-180-147-12 Sequence 12, Appli

78 166.5 8.4 134 4 US-09-540-245A-12 Sequence 12, Appli

79 166.5 8.4 134 4 US-09-540-153-12 Sequence 6, Appli

80 163 8.2 160 4 US-08-442-063A-60 Sequence 60, Appli

81 163 8.2 968 4 US-09-180-1439-6 Sequence 9, Appli

82 161.5 8.1 1012 2 US-08-057-315-4 Sequence 4, Appli

83 161.5 8.1 1025 2 US-08-057-315-4 Sequence 5, Appli

84 161.5 8.1 999 2 US-08-473-553A-5 Sequence 5, Appli

85 161 8.1 692 4 US-07-157-086-7 Sequence 73, Appli

86 160 8.0 863 2 US-08-166-271-2 Sequence 2, Appli

87 159 8.0 799 4 US-09-191-647-9 Sequence 9, Appli

88 159 8.0 735 4 US-09-540-153A-9 Sequence 9, Appli

89 159 8.0 735 4 US-09-140-047-12 Sequence 12, Appli

90 156 7.8 999 2 US-08-473-553A-5 Sequence 5, Appli

91 150 7.5 1023 2 US-08-166-271-2 Sequence 2, Appli

92 149 7.5 1023 2 US-08-166-271-2 Sequence 2, Appli

93 149 7.5 1023 2 US-08-166-271-2 Sequence 2, Appli

94 149 7.5 1023 2 US-08-166-271-2 Sequence 2, Appli

95 147.5 7.4 910 4 US-09-191-647-8 Sequence 8, Appli

96 145.5 7.3 1196 4 US-08-081-706-2 Sequence 2, Appli

97 143.5 7.2 155 3 US-09-191-647-8 Sequence 8, Appli

98 143.5 7.2 155 4 US-09-540-153-8 Sequence 8, Appli

99 143.5 7.2 155 4 US-09-540-153-8 Sequence 8, Appli

100 142.5 7.2 105 3 US-09-191-647-3 Sequence 3, Appli

101	7.2	105	4	US-09-540-245A-3	Sequence 3, Appli	174	100.5	456	US-08-795-395-2	
102	142.5	105	4	US-09-540-153-3	Sequence 3, Appli	175	99	1366	US-09-004-838-22	
103	140.5	7.1	390	3	US-08-460-576-2	Sequence 2, Appli	176	99	1890	US-09-004-838-88
104	140.5	7.1	695	1	US-08-487-886-2	Sequence 2, Appli	177	98.5	4.9	US-07-757-342D-5
105	140.5	7.1	695	3	US-08-482-855-2	Sequence 2, Appli	178	98.5	4.9	US-07-741-453A-59
106	140.5	7.1	695	4	US-08-74-986-2	GENERAL_INFORMA	179	98.5	4.9	US-07-741-453A-61
107	134.5	6.8	968	4	US-09-228-986-76	Sequence 76, Appli	180	98	4.9	US-08-724-303-4
108	134.5	6.8	644	4	US-08-866-757-2	Sequence 2, Appli	181	98	4.9	US-08-908-436-2
109	134.5	6.8	644	4	US-08-153-593-2	Sequence 2, Appli	182	98	4.9	US-09-561-561-18
110	6.4	740	1	US-08-257-073-5	Sequence 2, Appli	183	98	4.9	US-09-107-59B-7	
111	125.5	6.3	342	1	US-08-244-646-15	Sequence 15, Appli	184	98	4.9	US-09-227-721-18
112	125.5	6.3	342	1	US-08-592-936B-21	Sequence 21, Appli	185	98	4.9	US-07-741-453A-35
113	125.5	6.3	342	2	US-09-111-573-21	Sequence 21, Appli	186	98	4.9	US-08-786-305A-5
114	125.5	6.3	771	4	US-09-188-930-183	Sequence 183, Appli	187	98	4.9	US-08-741-104A-5
115	125	6.3	674	4	US-07-757-342D-10	Sequence 10, Appli	188	98	4.9	US-08-440-816A-5
116	125	6.3	699	4	US-07-757-342D-2	Sequence 2, Appli	189	98	4.9	US-09-117-381A-5
117	122	6.1	523	5	PCT-US91-09055-8	Sequence 3, Appli	190	97	4.9	US-08-442-063A-30
118	122	6.1	690	4	US-08-473-553A-3	Sequence 6, Appli	191	97	4.9	US-09-134-001C-4930
119	122	6.1	980	2	US-08-473-553A-6	Sequence 4, Appli	192	97	4.9	US-07-157-342D-8
120	121	6.1	985	4	US-08-473-553A-2	Sequence 2, Appli	193	97	4.9	US-07-757-342D-7
121	120.5	6.0	596	4	US-09-752-165-2	Sequence 17, Appli	194	97	4.9	US-09-134-001C-4588
122	118	5.9	227	1	US-08-244-646-17	Sequence 3, Appli	195	97	4.9	US-08-517-578-2
123	116.5	5.9	700	4	US-07-757-342D-3	Sequence 8, Appli	196	96.5	4.8	US-08-442-063A-30
124	115.5	5.8	196	5	PCT-US91-09055-8	Sequence 69, Appli	197	96	4.8	US-09-134-001C-4930
125	114	5.7	690	4	US-09-228-986-69	Sequence 4, Appli	198	95.5	4.8	US-08-930-996A-8
126	112.5	5.7	696	4	US-07-757-342D-4	Sequence 79, Appli	199	95.5	4.8	US-08-375-300-4
127	112	5.6	711	4	US-09-228-986-79	Sequence 4, Appli	200	95	4.8	US-09-177-301-4
128	110.5	5.5	327	1	US-08-238-163-4	Sequence 4, Appli	201	95	4.8	US-08-741-646-8
129	110.5	5.5	456	2	US-08-910-731-4	Sequence 8, Appli	202	95	4.8	US-09-177-301-4
130	110.5	5.5	456	2	US-08-910-731-8	Sequence 4, Appli	203	95	4.8	PCT-US5-16930-4
131	110	5.5	630	4	US-08-795-395-71	Sequence 71, Appli	204	95	4.8	US-08-175-300-2
132	109	5.5	630	4	US-09-228-986-71	Sequence 29, Appli	205	95	4.8	US-09-177-311-2
133	109	5.5	428	4	US-08-190-802A-29	Sequence 29, Appli	206	95	4.8	US-09-134-001C-235
134	109	5.5	428	4	US-08-477-346-29	Sequence 29, Appli	207	94.5	4.7	US-09-108-315-2
135	109	5.5	456	2	US-08-910-731-4	Sequence 29, Appli	208	94.5	4.7	US-08-746-883-6
136	109	5.5	456	2	US-08-795-395-71	Sequence 7, Appli	209	94.5	4.7	US-09-651-656-21
137	109	5.5	850	4	US-08-441-104A-7	Sequence 7, Appli	210	94.5	4.7	US-09-650-855-21
138	109	5.5	850	2	US-08-440-816A-7	Sequence 7, Appli	211	94	4.7	US-09-156-923-2
139	109	5.5	850	4	US-09-417-381A-7	Sequence 7, Appli	212	94	4.7	US-08-159-105B-4
140	108.5	5.4	1085	1	US-08-473-089-29	Sequence 28, Appli	213	94	4.7	US-08-286-846A-4
141	108.5	5.4	1085	4	US-08-487-072A-29	Sequence 28, Appli	214	94	4.7	US-08-457-880A-2
142	108.5	5.4	1085	4	US-09-345-294-28	Sequence 28, Appli	215	94	4.7	US-08-444-622A-4
143	107.5	5.4	707	4	US-09-228-986-80	Sequence 80, Appli	216	94	4.7	US-08-942-562-4
144	107	5.4	903	4	US-09-228-986-78	Sequence 78, Appli	217	94	4.7	US-09-156-923-2
145	106	5.4	825	1	US-07-912-952-2	Sequence 2, Appli	218	94	4.7	US-08-159-105B-2
146	105	5.3	1085	1	US-08-431-080-28	Sequence 28, Appli	219	94	4.7	US-08-286-846A-2
147	105	5.3	1085	2	US-08-938-534-28	Sequence 8, Appli	220	94	4.7	US-08-457-880A-2
148	105	5.3	1085	4	US-08-286-846A-8	Sequence 8, Appli	221	94	4.7	US-08-444-622A-2
149	105	5.3	612	3	US-08-444-622A-6	Sequence 8, Appli	222	94	4.7	US-08-942-562-2
150	105	5.3	612	3	US-08-942-562-8	Sequence 8, Appli	223	94	4.7	US-09-156-923-2
151	105	5.3	612	4	US-09-156-923-8	Sequence 8, Appli	224	94	4.7	US-09-245-381-43
152	105	5.3	839	2	US-08-359-705B-2	Sequence 6, Appli	225	94	4.7	US-09-207-359B-43
153	105	5.3	839	2	US-08-286-846A-6	Sequence 6, Appli	226	93.5	4.7	US-08-130-996A-4
154	105	5.3	839	3	US-08-444-622A-6	Sequence 6, Appli	227	93	4.7	US-09-228-386-74
155	105	5.3	839	3	US-08-444-622A-6	Sequence 6, Appli	228	93	4.7	US-09-022-383-2
156	105	5.3	839	3	US-08-942-562-6	Sequence 6, Appli	229	93	4.7	US-08-930-996A-7
157	105	5.3	839	4	US-09-156-923-6	Sequence 6, Appli	230	93	4.7	US-09-099-094-8
158	105	5.3	2391	4	US-09-150-741-2	Sequence 2, Appli	231	93	4.7	US-09-245-281-4
159	105	5.3	764	4	US-07-741-453A-60	Sequence 2, Appli	232	93	4.7	US-09-207-359B-8
160	104	5.2	620	3	US-09-126-646-2	Sequence 2, Appli	233	93	4.7	US-09-180-846A-2
161	104	5.2	620	4	US-09-421-491-2	Sequence 2, Appli	234	92.5	4.6	US-08-477-396A-4
162	103.5	5.2	723	4	US-09-434-408-2	Sequence 2, Appli	235	92.5	4.6	US-09-022-383-2
163	103	5.2	330	1	US-08-238-163-2	Sequence 2, Appli	236	92.5	4.6	US-08-123-70B-2
164	102.5	5.1	764	4	US-07-741-453A-54	Sequence 54, Appli	237	92.5	4.6	US-08-054-441-2
165	102	5.1	2391	4	US-09-150-741-2	Sequence 60, Appli	238	92	4.6	US-09-004-838-119
166	102	5.1	277	4	US-07-741-453A-58	Sequence 58, Appli	239	91.5	4.6	US-08-130-996A-2
167	102	5.1	285	3	US-08-482-085B-20	Sequence 20, Appli	240	91	4.6	US-08-431-080-26
168	102	5.1	461	2	US-08-910-731-6	Sequence 6, Appli	241	91	4.6	US-08-938-534-26
169	101	5.1	172	1	US-08-318-947A-11	Sequence 11, Appli	242	91	4.6	US-09-345-294-26
170	101	5.1	172	2	US-08-795-303-11	Sequence 11, Appli	243	91	4.6	US-09-004-838-40
171	101	5.1	172	3	US-09-358-580-14	Sequence 14, Appli	244	91	4.6	US-07-741-453A-56
172	101	5.1	655	4	US-09-228-986-70	Sequence 70, Appli	245	91	4.6	US-08-310-12A-108
173	100.5	5.0	456	2	US-08-910-731-2	Sequence 2, Appli	246	91	4.6	US-09-301-085-108

247	91	4.6	1143	PCT-US95-045589-108	Sequence 108, App	853	4	US-08-913-880C-17
248	91	4.6	1144	US-08-261-663A-2	Sequence 2, Appli	858	4	US-08-913-880C-16
249	91	4.6	1144	US-08-261-663A-4	Sequence 4, Appli	860	4	US-08-913-880C-15
250	91	4.6	1144	US-08-930-996A-9	Sequence 9, Appli	862	4	US-08-913-880C-14
251	91	4.6	1144	PCT-US95-0754A-2	Sequence 3, Appli	865	4	US-08-913-880C-13
252	91	4.6	1144	PCT-US95-07754A-4	Sequence 2, Appli	866	4	US-08-913-880C-12
253	91	4.6	1144	PCT-US95-04589-107	Sequence 4, Appli	874	4	US-08-913-880C-11
254	91	4.6	1209	PCT-US95-04589-107	Sequence 107, Appli	875	4	US-08-913-880C-10
255	91	4.6	1258	US-08-310-912A-107	Sequence 107, App	876	4	US-08-913-880C-1
256	91	4.6	1258	US-08-301-085-107	Sequence 107, App	877	4	US-08-913-880C-1
257	91	4.6	1279	US-09-724-517-2	Sequence 2, Appli	878	4	US-09-572-191-2
258	91	4.6	1279	US-09-641-807A-2	Sequence 2, Appli	879	4	US-09-573-262-2
265	90.5	4.5	1312	US-08-592-126-148	Sequence 148, App	880	4	US-09-177-162A-22
266	90.5	4.5	1312	US-08-687-100-51	Sequence 2, Appli	881	4	US-09-198-955A-2
267	90.5	4.5	1854	US-09-004-838-108	Sequence 10, Appli	882	4	US-09-004-838-92
268	90	4.5	4302	US-08-658-136-5	Sequence 4, Appli	883	4	US-09-004-838-93
269	90	4.5	645	US-08-592-126-144	Sequence 144, App	884	4	US-09-079-765-5
270	90	4.5	645	US-08-680-080-47	Sequence 47, App	885	4	US-09-196-525-5
271	89.5	4.5	659	US-09-228-986-75	Sequence 75, Appli	886	4	US-09-318-717-5
272	89	4.5	339	US-09-052-469-8	Sequence 6, Appli	887	4	US-09-73-219-2
276	88.5	4.4	659	US-09-592-126-148	Sequence 148, Appli	888	4	US-09-198-955A-2
277	88.5	4.4	376	US-09-564-954-2	Sequence 2, Appli	889	4	US-09-198-955A-2
278	88.5	4.4	138	US-09-191-647-4	Sequence 51, Appli	890	4	US-09-198-955A-2
273	89	4.5	138	US-09-540-454A-4	Sequence 108, Appli	891	4	US-09-198-955A-2
274	89	4.5	138	US-09-540-153-4	Sequence 4, Appli	892	4	US-09-198-955A-2
275	89	4.5	4302	US-08-658-136-5	Sequence 5, Appli	893	4	US-09-198-955A-2
276	88.5	4.4	883	US-08-460-751-2	Sequence 2, Appli	894	4	US-09-198-955A-2
282	88.5	4.4	909	US-09-310-912A-2	Sequence 2, Appli	895	4	US-09-286-845A-9
283	88.5	4.4	885	US-08-841-089-2	Sequence 2, Appli	896	4	US-09-286-845A-9
284	88	4.4	885	US-09-301-085-2	Sequence 2, Appli	897	4	US-09-286-845A-9
279	88.5	4.4	885	PCT-US95-04570-2	Sequence 2, Appli	898	4	US-09-286-845A-9
280	88.5	4.4	885	PCT-US95-04589-142	Sequence 2, Appli	899	4	US-09-286-845A-9
281	88.5	4.4	909	US-09-310-912A-2	Sequence 142, Appli	900	4	US-09-286-845A-9
282	88.5	4.4	909	US-09-301-085-2	Sequence 142, Appli	901	4	US-09-286-845A-9
283	88.5	4.4	909	PCT-US95-04589-142	Sequence 142, Appli	902	4	US-09-286-845A-9
284	88	4.4	885	US-09-538-871-4	Sequence 4, Appli	903	4	US-09-286-845A-9
285	88	4.4	292	US-09-538-871-2	Sequence 2, Appli	904	4	US-09-286-845A-9
286	88	4.4	652	US-08-559-896B-2	Sequence 2, Appli	905	4	US-09-286-845A-9
287	87.5	4.4	652	US-08-442-063A-54	Sequence 54, Appli	906	4	US-09-286-845A-9
288	87.5	4.4	196	US-09-383-586-11	Sequence 11, Appli	907	4	US-09-286-845A-9
289	87	4.4	312	US-08-477-451-29	Sequence 29, Appli	908	4	US-09-286-845A-9
290	87.5	4.4	376	US-08-983-502-11	Sequence 11, Appli	909	4	US-09-286-845A-9
291	87.5	4.4	376	PCT-US96-10521-11	Sequence 11, Appli	910	4	US-09-286-845A-9
292	87.5	4.4	1727	US-08-477-451-10	Sequence 10, Appli	911	4	US-09-286-845A-9
293	87	4.4	50	US-08-290-736C-9	Sequence 9, Appli	912	4	US-09-286-845A-9
294	87	4.4	73	US-08-290-736C-8	Sequence 8, Appli	913	4	US-09-286-845A-9
295	87	4.4	274	US-08-477-451-29	Sequence 39/4, Appli	914	4	US-09-286-845A-9
296	87	4.4	553	US-08-134-001C-3924	Sequence 10, Appli	915	4	US-09-286-845A-9
297	87	4.4	1435	US-08-568-459A-4	Sequence 4, Appli	916	4	US-09-286-845A-9
298	87	4.4	1435	US-08-487-826B-4	Sequence 4, Appli	917	4	US-09-286-845A-9
299	87	4.4	1435	US-09-210-288-4	Sequence 4, Appli	918	4	US-09-286-845A-9
300	86.5	4.3	1066	US-09-004-838-24	Sequence 24, Appli	919	4	US-09-286-845A-9
301	85.5	4.3	1130	US-08-519-547A-6	Sequence 6, Appli	920	4	US-09-286-845A-9
302	86.5	4.3	372	US-09-345-882-29	Sequence 29, Appli	921	4	US-09-286-845A-9
303	86.5	4.3	509	US-08-328-322-19	Sequence 90, Appli	922	4	US-09-286-845A-9
304	86	4.3	300	534094-6	Patent No. 5340934	923	4	US-09-286-845A-9
305	86	4.3	983	US-09-134-001C-3814	Sequence 4, Appli	924	4	US-09-286-845A-9
306	86	4.3	3169	US-09-453-702B-257	Sequence 25/7, Appli	925	4	US-09-286-845A-9
307	85.5	4.3	1066	US-09-134-001C-4707	Sequence 47/7, Appli	926	4	US-09-286-845A-9
308	85.5	4.3	372	US-09-324-455-2	Sequence 2, Appli	927	4	US-09-286-845A-9
309	85.5	4.3	509	US-09-004-838-90	Sequence 19, Appli	928	4	US-09-286-845A-9
310	85.5	4.3	548	US-09-601-691-2	Sequence 2, Appli	929	4	US-09-286-845A-9
311	85.5	4.3	548	US-09-210-091-4	Sequence 2, Appli	930	4	US-09-286-845A-9
312	85.5	4.3	548	US-09-398-395A-52	Sequence 52, Appli	931	4	US-09-286-845A-9
313	85.5	4.3	548	US-09-514-502-3	Sequence 3, Appli	932	4	US-09-286-845A-9
314	85.5	4.3	1938	US-09-514-302-2	Sequence 2, Appli	933	4	US-09-286-845A-9
315	85	4.3	714	US-08-990-114-3	Sequence 3, Appli	934	4	US-09-286-845A-9
316	85	4.3	714	US-09-241-533-3	Sequence 3, Appli	935	4	US-09-286-845A-9
317	85	4.3	764	US-09-370-838-67	Sequence 3, Appli	936	4	US-09-286-845A-9
318	85	4.3	781	US-08-337-134D-2	Sequence 2, Appli	937	4	US-09-286-845A-9
319	85	4.3	781	US-09-114-637-2	Sequence 2, Appli	938	4	US-09-286-845A-9

393	83	4.2	591	2	US-08-937-540-4	Sequence 4, Appli	466	78.5	3.9	251	2	US-08-766-738-1
394	83	4.2	591	4	US-09-398-395A-24	Sequence 24, Appli	467	78.5	3.9	251	2	US-08-766-738-3
395	83	4.2	666	4	US-09-228-986-68	Sequence 68, Appli	468	78.5	3.9	251	4	US-09-262-610-3
396	83	4.2	693	1	US-08-463-620-11	Sequence 11, Appli	469	78.5	3.9	251	4	US-08-262-610-3
397	83	4.2	693	2	US-08-224-917-11	Sequence 11, Appli	470	78.5	3.9	394	1	US-08-672-514-3
398	83	4.2	693	2	US-08-914-853-11	Sequence 11, Appli	471	78.5	3.9	394	1	US-08-672-514-3
399	83	4.2	693	5	PCT-US95-0393A-11	Sequence 11, Appli	472	78.5	3.9	563	4	US-08-134-001C-3172
400	83	4.2	887	1	US-07-106-3	Sequence 3, Appli	473	78.5	3.9	3218	1	US-08-764-100-27
401	83	4.2	976	4	US-09-104-324B-4	Sequence 4, Appli	474	78.5	3.9	3696	4	US-09-134-001C-5080
402	82.5	4.1	411	2	US-08-741-134-6	Sequence 6, Appli	475	78	3.9	302	1	US-07-640-029-6
403	82.5	4.1	506	2	US-08-849-480A-5	Sequence 5, Appli	476	78	3.9	302	1	US-08-672-807B-8
404	82	4.1	153	4	US-09-228-986-103	Sequence 103, Appli	477	78	3.9	302	1	US-08-441-944A-8
405	82	4.1	515	3	US-08-930-996A-12	Sequence 12, Appli	478	78	3.9	302	4	US-08-439-992A-6
406	82	4.1	567	4	US-09-001C-3762	Sequence 31/2, Appli	479	78	3.9	305	4	PCT-US94-0225A-4
407	82	4.1	990	2	US-08-392-625-20	Sequence 20, Appli	480	78	3.9	383	2	US-08-484-575A-4
408	82	4.1	990	2	US-08-466-961A-20	Sequence 20, Appli	481	78	3.9	383	3	US-08-477-459-4
409	81.5	4.1	249	1	US-08-466-603-2	Sequence 2, Appli	482	78	3.9	383	3	US-08-479-869-4
410	81.5	4.1	249	1	US-08-314-503A-2	Sequence 2, Appli	483	78	3.9	383	4	US-08-486-414-4
411	81.5	4.1	249	1	US-08-468-066-2	Sequence 2, Appli	484	78	3.9	383	5	PCT-US94-01826A-4
412	81.5	4.1	249	2	US-08-466-717-2	Sequence 2, Appli	485	78	3.9	383	5	PCT-US94-0225A-4
413	81.5	4.1	249	2	US-08-766-738-4	Sequence 4, Appli	486	78	3.9	412	2	US-08-741-134-2
414	81.5	4.1	249	3	US-08-466-743-2	Sequence 2, Appli	487	78	3.9	530	4	US-08-975-762-3
415	81.5	4.1	249	3	US-09-262-610-4	Sequence 4, Appli	488	78	3.9	510	4	US-08-295-028-73
416	81.5	4.1	249	5	PCT-US95-12414-2	Sequence 2, Appli	489	78	3.9	530	4	US-09-106-582-3
417	81.5	4.1	282	1	US-07-712-476A-5	Sequence 5, Appli	490	78	3.9	540	1	US-08-286-325A-2
418	81.5	4.1	296	1	US-07-12-476A-1	Sequence 1, Appli	491	78	3.9	590	3	US-08-893-852A-4
419	81.5	4.1	448	4	US-09-134-001C-4146	Sequence 14, Appli	492	78	3.9	590	4	US-08-821-818-2
420	81.5	4.1	990	2	US-08-645-193B-15	Sequence 15, Appli	493	78	3.9	590	4	US-08-975-762-64
421	80.5	4.0	247	4	US-09-228-986-105	Sequence 105, Appli	494	78	3.9	590	4	US-08-295-028-64
422	80.5	4.0	337	3	US-09-190-965-1	Sequence 1, Appli	495	78	3.9	590	4	US-09-106-582-4
423	80.5	4.0	337	4	US-09-470-253-1	Sequence 1, Appli	496	78	3.9	701	4	US-09-132-028-2
424	80.5	4.0	496	4	US-09-204-917	Sequence 2, Appli	497	78	3.9	733	1	US-07-640-029-2
425	80.5	4.0	749	1	US-08-046-508-2	Sequence 2, Appli	498	78	3.9	733	1	US-07-921-807B-6
426	80.5	4.0	1604	4	US-09-004-838-95	Sequence 95, Appli	499	78	3.9	733	1	US-08-441-944A-6
427	80.5	4.0	1786	4	US-08-973-462-8	Sequence 8, Appli	500	78	3.9	733	4	US-08-439-992A-4
428	80.5	4.0	10182	4	US-09-134-001C-3159	Sequence 31/59, Appli						
429	80	4.0	563	4	US-09-134-001C-4800	Sequence 4800, Appli						
430	80	4.0	574	6	522/254-2	Patent No. 5232354						
431	80	4.0	896	1	US-08-095-737-2	Sequence 2, Appli						
432	80	4.0	896	1	US-08-480-145-2	Sequence 2, Appli						
433	80	4.0	896	2	US-08-477-389-2	Sequence 2, Appli						
434	80	4.0	1038	4	US-09-541-782-4	Sequence 4, Appli						
435	80	4.0	1038	4	US-09-723-820-4	Sequence 4, Appli						
436	80	4.0	1235	1	US-08-104-101A-2	Sequence 14, Appli						
437	80	4.0	1420	2	US-08-540-804-14	Sequence 14, Appli						
438	80	4.0	1420	2	US-08-218-265-14	Sequence 14, Appli						
439	80	4.0	1420	3	US-08-521-872-14	Sequence 14, Appli						
440	80	4.0	1420	4	US-08-590-399-14	Sequence 14, Appli						
441	80	4.0	1507	3	US-08-928-329-5	Sequence 5, Appli						
442	79.5	4.0	677	3	US-08-480-640A-115	Sequence 115, Appli						
443	79.5	4.0	677	3	US-08-452-239-22	Sequence 22, Appli						
444	79.5	4.0	677	4	US-09-004-838-50	Sequence 50, Appli						
445	79.5	4.0	677	4	US-08-688-968C-58	Sequence 58, Appli						
446	79.5	4.0	677	4	US-08-914-824-3	Sequence 193, Appli						
447	79.5	4.0	677	4	US-08-488-237A-115	Sequence 4053, Appli						
448	79.5	4.0	677	4	US-08-480-640A-115	Sequence 115, Appli						
449	79.5	4.0	677	3	US-08-452-239-22	Sequence 115, Appli						
450	79.5	4.0	677	3	US-08-298-802-115	Sequence 115, Appli						
451	79.5	4.0	677	4	US-08-688-968C-193	Sequence 115, Appli						
452	79.5	4.0	677	4	US-08-914-824-3	Sequence 115, Appli						
453	79.5	4.0	677	4	US-08-488-237A-115	Sequence 115, Appli						
454	79.5	4.0	677	4	US-08-480-640A-193	Sequence 115, Appli						
455	79.5	4.0	677	4	US-08-375-992A-115	Sequence 115, Appli						
456	79.5	4.0	677	4	US-08-298-802-115	Sequence 115, Appli						
457	79.5	4.0	677	4	US-08-688-968C-193	Sequence 115, Appli						
458	79.5	4.0	677	4	US-08-914-824-3	Sequence 115, Appli						
459	79	4.0	255	3	US-09-126-657-9	Sequence 9, Appli						
460	79	4.0	311	4	US-09-305-16	Sequence 16, Appli						
461	79	4.0	470	4	US-09-004-838-44	Sequence 44, Appli						
462	79	4.0	527	3	US-08-923-958-2	Sequence 2, Appli						
463	79	4.0	527	4	US-08-485-553-2	Sequence 2, Appli						
464	79	4.0	759	4	US-09-199-637A-93	Sequence 170, Appli						
465	79	4.0	811	4	US-09-199-637A-93	Sequence 93, Appli						

ALIGNMENT S

```

5: sp_human:*
6: sp_invertebrate:*
7: sp_mammal:*
8: sp_mcbe:*
9: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rabbit:*
11: sp_rat:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriophage:*
17: sp_archaea:*

Pred No is the number of records recorded.
      RESULT 1
      Q9DE04 ID Q9DE04; PRELIMINARY;
      AC Q9DE04; 370 AA.
      DT 01-MAR-2001 (TREMBrel. 16, Created)
      DT 01-MAR-2001 (TREMBrel. 16, Last sequence update)
      DT 01-JUN-2001 (TREMBrel. 17, Last annotation update)
      DE Biglycan-like protein 3.
      GN BG13.
      OS Oreochromis niloticus ('Nile tilapia') (Tilapia nilotica).
      OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
      OC Actinopterygii; Neopterygii; Teleostei; Buteleosteii; Neoteleosteii;
      OC Acanthorhormia; Acanthopterygii; Perciformes; Labroidei;
      OC Cichlidae; Oreochromis.

```

score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMATES

result	No.	Score	Query Match	Length	DB ID	Description
	1	1118.5	56..1	370	13 Q9DE04	O9de04 oreochromis
	2	968	48..6	187	4 Q9NXP3	O9nxp3 homo sapien
	3	9	47..7	359	13 Q9DE03	O9de03 oreochromis
	4	917	46..0	410	13 Q9DDZ7	O9ddz7 petromyzon
	5	869.5	45..9	310	13 Q9DDZ8	O9ddz8 petromyzon
	6	865..5	43..4	347	13 Q9DE00	O9de00 petromyzon
	7	863..5	43..4	388	13 Q9DDZ9	O9ddz9 petromyzon
	8	863..5	32..1	224	13 Q9DE01	O9de01 brachydanio
	9	520	26..1	108	11 Q9CTL6	Q9ctl6 mus musculus
	10	397.5	20..0	699	4 Q94769	O94769 homo sapien
	11	363	18..2	674	4 Q8WVA2	Q8wva2 homo sapien
	12	354	17..8	120	13 Q9DE02	O9de02 brachydanio
	13	340..5	17..1	796	11 Q9WV1	Q9wv1 ratmus norvegicus
	14	336	16..9	1525	4 Q9Y5Q7	Q9y5q7 homo sapien
	15	336	16..9	1529	4 Q9A813	Q9a813 homo sapien
	16	335	16..8	1512	13 Q9DE36	O9de36 brachydanio
					SEQUENCE 370 AA; 42177 MW;	BCD0675694ECA2B7 CRC64;
					SQ	
					Query Match	56..1%; Score 1118..5; DB 13; Length 370;
					Best Local Similarity	56..7%; Pred. No. 7e-72;
					Matches 219;	Mismatches 88; Indels 27; Caps
					Conservative	52;
					QY	1 MKEVYLFLALCSAKKPFPSPSHIA-LKN-----MMLKDMEIDTDDDDDDDDDE 51

RESULT	2	NINP3	Q9NXP3;	PRELIMINARY:	PRT;	187 AA.
		0	0	(TREMBLrel. 15 Created)		
		1	0	01-OCT-2000 (TREMBLrel. 15, Last sequence update 01-OCT-2000 (TREMBLrel. 15, Last annotation update 01-JUN-2002 (TREMBLrel. 21, Last annotation update CDNA FLJ20129 firs, clone CCL06190.		
		2	0	Homo sapiens (Human).		
		3	0	Eukaryota; Metazaca; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Hominidae; Hominina; Hominini; Homininae; H. Sapiens.		

SEQUENCE FROM N.A.  
TISSUE-COLON;  
Kawabata A., Hikiji T., Kobatake N., Inagaki H.,  
Okitani R., Ota T., Suzuki Y., Obayashi M., Nishimura T., Nakamura Y., Isgai T., Sugano S.;  
NEDMO Human DNA sequencing project;  
Submitted (FEBR-2000) to the EMBL/GenBank/DBJ/

InterpO; IPR01611; LRR.  
Pfam; PF00560; LRR; 3.  
SEQUENCE 187 AA; 21074 MW; BDBEE37373CA3DBE CRC64;

102 MINIMUM ENERGY DYNAMIC EQUILIBRIUM STATE OF A POLYMER 25  
103 MINIMUM ENERGY DYNAMIC EQUILIBRIUM STATE OF A POLYMER 25

1 MNAIHVLENSANPDLNGLEPGAEFGTVFHIRIAEAKLTSVKGLPPTLIELHDYNI 60  
253 STVLEDFKRYKEQLRQLGICNNKTDIENGSLANIPVRBLHENNKLKKPSGLPELK 31

b	b	311.3	LQI1-FLHSNSIARYGVNDFCPTVKMKKSIYSAISIUNNPVKYWEMOPATRCVLSRMSI 311.4	37
b	b	61	STVLEDFRYKEFQLGLGNKTDIENGSLANIPVRTEHNNKLKIPSGLELK Y 62	

1.2.1 EQUATIONS OF MOTION FOR A Rigid BODY

GN BGL2	RL "Biglycan-like extracellular matrix genes of agnathans and teleosts." ; J. Mol. Evol. 51:363-373 (2000).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;	RT EMBL; AF247827; AAC4162.; -.
OC Petromyzontiformes; Petromyzontidae; Petromyzon.	DR InterPro; IPR001611; LRR.
OX NCBI_TaxID=7757;	DR InterPro; IPR000372; LRR_Nterm.
RN [1]	DR InterPro; IPR003592; LRR_Out.
RP SEQUENCE FROM N.A.	DR InterPro; IPR003591; LRR_TYP.
RA MEDLINE=20496956; PubMed=11040287;	DR Pfam; PF000560; LRR; 7.
RT Shintani S., Sato A., Toyosawa S., O'hugain C., Klein J.; R.L. J. Mol. Evol. 51:363-373 (2000).	DR Pfam; PF01462; LRRNT; 1.
DR EMBL; AF247828; AAC4163.1.; -.	DR SMART; SM00370; LRR; 4.
DR InterPro; IPR001611; LRR.	DR SMART; SM00013; LRRNT; 1.
DR InterPro; IPR000372; LRR_Nterm.	DR SMART; SM00369; LRR_TYP; 8.
DR InterPro; IPR003592; LRR_Out.	FT NON_TER 1 1
DR InterPro; IPR003591; LRR_TYP.	SEQUENCE 310 AA; 34572 MW; 6FB8E05B14339EF5 CRC64.
DR Pfan; PF000560; LRR; 7.	Query Match 45.9%; Score 915; DB 13; Length 310;
DR Pfan; PF01462; LRRNT; 1.	Best Local Similarity 56.2%; Pred. No. 1.7e-57;
DR SMART; SM00370; LRR; 4.	Matches 172; Conservative 51; Mismatches 81; Indels 2; Gaps 2;
DR SMART; SM00013; LRRNT; 1.	QY 74 CPFGCQCYSRVHCSDDLGTSVPTNPFLDQLLQNNKIKEKENDFKGLTSLYGLILN 133
DR SMART; SM00369; LRR_TYP; 8.	Db 3 CPFGCQCSARVQQCDLGVSPOAIPKARLLQNNKITEKDDPKGLNLKYALYLV 62
SEQUENCE 410 AA; 44671 MW; 3EC96E490BCCETA CRC64;	QY 134 NNKLTKTHPKAELTTPKKLRLYLSHNOLSEPIPLPKSLAAELRHENKVKKIKDTPKGMI 193
Query Match 46.0%; Score 917; DB 13; Length 410;	Db 63 NNLSKVKHPKAFAPLSSDPLKLYLISHNOLTEPVPGSMPSLAVELRHNKKIKERDAFSGM 122
Best Local Similarity 46.5%; Pred. No. 1.7e-57;	QY 194 NALHYLEMSANPLDNNGTIPGAGFEGV-TVFHIRIAEKLTISVPKGGLPDTLELHDYNKI 252
Matches 192; Conservative 58; Mismatches 109; Indels 54; Gaps 7;	Db 123 KRLHALEMCGNPLQSTGVGAEGSLERLVYVRSYDSKAKRKPDLPSIQELHLHNQI 182
QY 54 SLFPTREPR-----SHFFPDLPF-----MCPRGCQCYSRVHV 86	QY 253 STVELEDFKRYKELQRLGQNLKNTDINGSLANTIPREHLENNKLKIPSGLPELK 312
Db 56 SWAPSAPPKVKGDRSKATAGKOPGQAATPKSLPPPPPPDASCPCGQCSARVQ 115	Db 183 TALEOFLIRPLIHRGLSYNQIKVIONGSLETCPHRELHDNSVLTQVPLGLAFIKH 242
QY 87 CSDLGLTSVPTNIPFDTRMLDQNLKMDNDTDDDDDDDEDDN 53	QY 313 LQIPLHNSIARYGVNDFCPTVPKMKKSLYSATSLSFNNPVKWMQPATFRCVLSRMSV 372
Db 9 LLVCALSSSPSSSSVATATSSKPF----AQROFFTDEMAADAVERSGDD----- 55	Db 243 LQVYVLHSNKIAAVKSDDPCSKGASPKRVLYSGSLFDNPVNWDPVSAFRCVASRAV 302
QY 54 SLFPTREPR-----SHFFPDLPF-----MCPRGCQCYSRVHV 86	QY 373 QLG-NF 377
Db 56 SWAPSAPPKVKGDRSKATAGKOPGQAATPKSLPPPPPPDASCPCGQCSARVQ 115	Db 303 QFSQNF 308
QY 87 CSDLGLTSVPTNIPFDTRMLDQNLKMDNDTDDDDDDDEDDN 53	RESULT 6
Db 116 CSDPLGLSVVPQKIPDKARLLQNLKNTDINGSLANTIPREHLENNKLKIPSGLPELK 146	Q9DE00 PRELIMINARY; PRIT; 347 AA.
QY 147 TPKKLRLYLHSNQNLSTIPLNPKSLAAELRHENKVKTDRFGMNAJHLLEMANSPL 206	AC Q9DE00; PRELIMINARY;
Db 176 PLSSLDKLYISINQTLTEPVGSMPSLVELTHENNNIKPKDAGSGMKRHLAENGGNPL 235	AC Q9DE00; PRELIMINARY;
QY 207 DNGIEPGLAFEGV-TVFHIRIAEKLTISVPKGGLPDTLELHDYNKISTVELEDFKRYKE 265	AC Q9DE00; PRELIMINARY;
Db 236 QSTGIEVGAEGLERIVYVRSYDSKAKRIPDKLPSIQEHLHQITAEFOEDLIRYPL 295	AC Q9DE00; PRELIMINARY;
QY 266 LORGLGNNKNTDINGSLANTIPREHLENNKLKIPSGLPELKYLQIITLEHNSSTAR 325	AC Q9DE00; PRELIMINARY;
Db 296 IHLRLGSLNQKIVQVNGSLETCPHRELHDNSVLTQVPEGLAFKLHQLqVYVLSHSNKLAA 355	AC Q9DE00; PRELIMINARY;
DE Biglycan-like protein 2 (Fragment).	AC Q9DE00; PRELIMINARY;
GN BGL2.	AC Q9DE00; PRELIMINARY;
OS Petromyzon marinus (Sea lamprey).	AC Q9DE00; PRELIMINARY;
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;	AC Q9DE00; PRELIMINARY;
OC Petromyzontiformes; Petromyzontidae; Petromyzon.	AC Q9DE00; PRELIMINARY;
OC NCBI_TaxID=7757;	AC Q9DE00; PRELIMINARY;
RN [1]	RN SEQUENCE FROM N.A.
RP SEQUENCE FROM N.A.	RT MEDLINE=20496956; PubMed=11040287;
RX Shintani S., Sato A., Toyosawa S., O'hugain C., Klein J.;	DR InterPro; IPR001611; LRR.
RA 01-MAR-2001 (TREMBLrel. 16, Created)	DR InterPro; IPR00372; LRR_Nterm.
RA 01-MAR-2001 (TREMBLrel. 16, Last sequence update)	DR InterPro; IPR003592; LRR_Out.
RA 01-JUN-2002 (TREMBLrel. 21, Last annotation update)	DR InterPro; IPR003591; LRR_TYP.
DE Biglycan-like protein 1 (Fragment).	DR Pfam; PF000560; LRR; 8.
GN Petromyzon marinus (Sea lamprey).	DR Pfam; PF01462; LRRNT; 1.
OS Petromyzontiformes; Chordata; Craniata; Vertebrata; Hyperoartia;	DR SMART; SM00370; LRR; 4.
OC Petromyzontidae; Petromyzontidae; Petromyzon.	DR SMART; SM00013; LRRNT; 1.
OC NCBI_TaxID=7757;	DR SMART; SM00369; LRR_TYP; 6.

Wed May 28 09:45:57 2003

RESULT 9		SEQUENCE FROM N.A.	
Q9CTU6	PRELIMINARY;	PRT;	108 AA.
ID			
AC			
DT	01-JUN-2001 (TREMBLrel. 17, Created)		
DT	01-JUN-2001 (TREMBLrel. 17, Last sequence update)		
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)		
DE	4631401G09Rik protein (fragment).		
GN	ASPN OR 4631401G09Rik		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognath; Muridae; Murinae; Mus.		
NCBI_TaxID	10090; [1]		
RN			
RP	SEQUENCE FROM N.A.		
RX	Medline=99009324; PubMed=9790758;		
RA	Nishii J.; Tanaka T.; Nakamura Y.;		
RT	"Identification of a novel gene (ECM2) encoding a putative extracellular matrix protein expressed predominantly in adipose and female specific tissues and its chromosomal localization to 9q22.3.";		
RL	Genomics 52:378-381 (1998).		
DR	EMBL: AB011792; BAA31958.1; -.		
DR	InterPro; IPR01611; LRR.		
DR	InterPro; IPR003592; LRR_out.		
DR	InterPro; IPR01007; VWF_C.		
PFam	PF00356; LRR; 11.		
PFam	PF00093; vwc; 1.		
PRINTS	PRO0019; LEURICH_RPT.		
SMART	SM00370; LRR; 4.		
DR	PROSITE; SM00214; vwc; 1.		
DR	PROSITE; PS01208; vwc; UNKNOWN_1.		
KW	Matrix protein.		
SEQUENCE	69 AA; 79789 MW; E44E76A40A5C2742 CRC64; [1]		
Query Match	20.0%; Score: 397.5; DB 4; Length: 699;		
Best Local Similarity	24.9%; Pred. No. 3.1e-20;		
Matches	108; Conservative: 77; Mismatches: 148; Indels: 101; Gaps: 8;		
Qy	32 KDMEDDDDDDDDDDDDEDNSL-----FPTRPFRSHFPFDLFP-----MCPEGC 78		
Db	270 EEEDEEEGEEGEDEEDEDPVGRGDMFMRMPSPSP-----LPAPPRTGTLRPLSGC 320		
Qy	79 QCYSRVYHCSLDGLTSVPTNTPFDTRMLDQNKKIKKEIKENDFKGLTSYGLILNNKKLT 138		
Db	321 SLSYRTISCNAMLTQIPPLTAQITSLTEGTNSIASIDPEAFNGLPNLERLDSKNNT 380		
Qy	139 K- <i>IHPKAFLTTKKLRLYLPSHNOLSEIPINLPKSLAENRTHENKVKKIQRDTPKGMNA</i> 196		
Db	381 SSGIGKAKFLKLKLMRNLNDGNNLQIQPSOLSPLEELKVNNLQADEEISLDLNLG 440		
Qy	197 HVLEMANSPLDONGIEPGAFEGV-TVFHRIIAEKALTSYSPKGFLPTLLHDYNKISTV 255		
Db	441 VTLLEGNIUNLSEANVNPLAKPLKSLAYLRIGKNNKFRIPQGLGSIEAYLENQIEET 500		
Qy	256 ELEDFKRYKEYQLRGLGNNKI-----TDIENGSLNIP----- 288		
Db	501 TEICFNHTRKINVLYRNKIEENRIAPLAWINQENLESIDLSSYKLVPSYLPKSLLH 560		
Qy	289 ----- 300		
Db	561 LVLLGNGQTERTPGYVFGHMEPGLEYLFSNKLAADDGMDRVSFYGAHSRELFLDHNL 620		
Qy	301 KRIPSGLPIELKYLQITFLHNSIARVGNDVFCPTVPKMKKSLYSATISLFPVWEMQP 360		
Db	621 KSIPOCIQEMKALHFIRLNINKIRNLIPPEIC-NAEEDDSNLIEHLHNNYKIREIPS 679		
Qy	361 ATFCRVLSRMSVOL 374		
Db	680 YTFSCTRSYSSIVL 693		
RESULT 10			
Q9WVA2	PRELIMINARY;	PRT;	674 AA.
ID	Q9WVA2;		
AC	Q9WVA2;		
DT	01-MAR-2002 (TREMBLrel. 20, Created)		
DT	01-MAR-2002 (TREMBLrel. 20, Last sequence update)		
DE	Hypothetical 74.1 kDa protein.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
OC	NCBI_TaxID=9606; [1]		
OX	NCBI_TaxID=9606;		
RN			
RP	SEQUENCE FROM N.A.		
RC	TISSUE=BRAIN;		
RA	Strausberg R.;		

RL	Submitted (DEC-2001) to the EMBL/GenBank/DDBJ databases.	DR	SMART; SM00370; LRR; 3.
DR	EMBL: BC018370; AAH18370_1; -.	DR	SMART; SM00369; LRR_TYP; 4.
DR	InterPro; IPR003961; FN_ITT.	FT	NON_TER
DR	InterPro; IPR001611; LRR.	FT	NON_TER
DR	InterPro; IPR000483; LRR_Cterm.	FT	NON_TER
DR	InterPro; IPR000372; LRR_Nterm.	FT	NON_TER
DR	InterPro; IPR00592; LRR_out.	FT	NON_TER
DR	InterPro; IPR003591; LRR_TYP.	FT	NON_TER
DR	InterPro; IPR001211; phospholipaseA2.	FT	NON_TER
PFam;	PF00041; fn3; 1.	FT	NON_TER
DR	PFam; PF00560; LRR; 7.	FT	NON_TER
PFam;	PF01463; LRCT; 1.	FT	NON_TER
PFam;	PF01465; LRNT; 1.	FT	NON_TER
DR	PRINTS; PR00019; LEURICHRPT.	FT	NON_TER
SMART;	SM00082; LRRCT; 1.	FT	NON_TER
SMART;	SM00369; LRR_TYP; 1.	FT	NON_TER
DR	PROSITE; PS00119; PA2_ASPI; UNKNOWN_1.	FT	NON_TER
KW	Hypothetical protein.	FT	NON_TER
SEQUENCE	674 AA; 74087 MW;	FT	NON_TER
SEQUENCE	674 AA; 38AB53F7243166CC CRC64;	FT	NON_TER
Query Match	18.2%; Score 363.5; DB 4; Length 674;	FT	NON_TER
Best Local Similarity	32.7%; Pred. No. 7.8e-18;	FT	NON_TER
Matches	51; Mismatches 105; Indels 33; Gaps 7;	FT	NON_TER
Qy	74 CPFGCQCYSRVVHCSDLGLTSVPPTNIPFDTRMLDQNNTKEIKENDFKGLTSLYGLILN 133	FT	NON_TER
Db	54 CPSVCRCDNGFIYCNDRGLTSPADPDATTLYLQNNQI-----N 94	FT	NON_TER
Qy	134 NNKLTTKPAFLTTLKLRLPQLSHNOLSEIPLPNPKSLAELRTHENVKKICKDTFKGM 193	FT	NON_TER
Db	95 NAGI---PODLKTRKVNOVIVLYNDLEFPNLPRSRELQDNVVTRLARDISLARI 150	FT	NON_TER
Qy	194 NALHYLEMSANPLDONGTNGIEPGAF-EVGTVFHITRAEAKLTSVPKGGLPPTLLEHLHDYNKI 252	FT	NON_TER
Db	151 PLEKLHDDNSVSTVSIEEDAFSKQLKLFLSRNHSSISGLPLHTLEERLDNRI 210	FT	NON_TER
Qy	253 STEVEDEFKRYKYLORLGIGNNKTD-LENGSLANIPRVEFHLENNKLKKPSGELPEL 310	FT	NON_TER
Db	211 STIPLHAFKGLNLSLRLVLDGNLLANORLADEDSRQLNLTELSLVRVNSLAAPLNLP- 269	FT	NON_TER
Qy	311 KLYQTLTFLHSNSIARYGVNDFCPVPMKMKSLYSAISLENN 351	FT	NON_TER
Db	270 AHLQKLYLQDNAISHIPYN---TIAKMR--LERLDSNN 304	FT	NON_TER
RESULT 12			
ID	O9DE02	PRT;	120 AA.
AC	O9DE02	PRELIMINARY;	PRT;
DR	01-MAR-2001 (TREMBLrel. 16, Created)	FT	NON_TER
DR	01-MAR-2001 (TREMBLrel. 16, Last sequence update)	FT	NON_TER
DR	01-DEC-2001 (TREMBLrel. 19, Last annotation update)	FT	NON_TER
GN	BGL3 OR BGL3.	FT	NON_TER
OS	Brachydanio rerio (zebra fish) (zebra danio).	FT	NON_TER
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;	FT	NON_TER
OC	Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;	FT	NON_TER
OC	Cyprinidae; Danio.	FT	NON_TER
OX	NCBI_TaxID=7955;	FT	NON_TER
RN	[1]	FT	NON_TER
RP	SEQUENCE FROM N.A.	FT	NON_TER
RX	MEDLINE=20496956; PubMed=11040287;	FT	NON_TER
RA	Shintani S., Satoh A., Toyosawa S., O'hugain C., Klein J.;	FT	NON_TER
RA	J. MOJ. Evol. 51:363-373 (2000).	FT	NON_TER
DR	EMBL; AF247823; AGI0158.;	FT	NON_TER
DR	ZFIN; ZDB-GENE-010131-5; b913.	FT	NON_TER
DR	InterPro; IPR001611; LRR.	FT	NON_TER
DR	InterPro; IPR00592; LRR_out.	FT	NON_TER
DR	InterPro; IPR003591; LRR_TYP.	FT	NON_TER
PFam;	PF00560; LRR; 4.	FT	NON_TER
Qy	213 -FEGYTVVFH 223	FT	NON_TER

Protein BLAST Results											
Query	Subject	Length	Score	E-value	Bit Score	Expect	Identity	Mismatches	Gaps	Similarity	Details
456 TSGARTCSPRLANKRIGQIKSKKFRCSGTEDYRSKLSDGCFADLACPERCREGTV--	SMART; SMD0001; EGF-like; 7.	513	DR SMART; SMD0001; EGF-like; 7.	DR SMART; SMD00274; FOLN; 2.							
224 IRTAEAKLTSPVPGGLPPPLLEHLDYNIK1STVELED-FKRYKEKLQLRGIGNNK1TIDENG	SMART; SMD00282; LamC; 1.	282	DR SMART; SMD00282; LamC; 1.	DR SMART; SMD00370; LRR; 6.							
514 -DSNQKLANKIPIPHIQYTAELFLNNNFFETVLATG1KKLQLRK1INSNW1ITD1EFG	SMART; SMD0082; LRRCT; 4.	572	DR SMART; SMD0082; LRRCT; 4.	DR SMART; SMD0013; LRRNT; 4.							
283 SLANIPRVRKNEILTSNRLENVQKMFKGLESMTL--MLRSRISCVGNDSLSTGLGSVR	DR PROSITE; PS0010; ASX_HYDROXYL; UNKNOWN_2.	338	DR PROSITE; PS0010; ASX_HYDROXYL; UNKNOWN_2.	DR PROSITE; PS01165; CTCK_1; UNKNOWN_1.							
573 AFPGASGVNEILTSNRLENVQKMFKGLESMTL--MLRSRISCVGNDSLSTGLGSVR	DR PROSITE; PS01255; CTCK_2; 1.	629	DR PROSITE; PS00122; EGF_1; UNKNOWN_9.	DR PROSITE; PS01166; EGF_2; 7.							
339 KKSLY-----SAISLFNNPQKYWENQPATERC-----VLSRMS	DR PROSITE; PS01187; EGF_CA; 2.	371	DR PROSITE; PS01187; EGF_CA; 2.	DR PROSITE; PS01187; EGF_CA; 2.							
630 LLSDYDNOITTVAPGAFGTLHSISLSTNLJANP-----FNCNCHLAWLGEWLRKK	KW Calcium-binding; EGF-like domain; Glycoprotein; Repeat; SQ SEQUENCE 1525 AA: 169394 MW; 8A81CDBE34EF06A73 CRC64;	679									
372 VQIGN 376	Query Match Score: 336; DB 4; Length: 1525;		16.9%; Score: 336; DB 4;	Pred No: 1 9e-15;							
680 IVNGN 684	Best Local Similarity: 26.5%; Pred No: 1 9e-15; Mismatches: 61; Gaps: 120; Details: 110; Gaps: 12		16.9%; Score: 336; DB 4;	Pred No: 1 9e-15;							
Q9Y5Q7 PRELIMINARY; PRT: 1525 AA.	Query 74 CPFGCOCYSRVVHCSDLGLTSVPTNPFDT-----RMLDL 108		16.9%; Score: 336; DB 4;	Pred No: 1 9e-15;							
Q9Y5Q7; 09Y5Q7; AC_09Y5Q7; DT_01-NOV-1999 (TREMBLrel. 12, Created) DT_01-NOV-1999 (TREMBLrel. 12, Last sequence update) DT_01-JUN-2002 (TREMBLrel. 21, Last annotation update) DE_SLT2.	Db 277 CPAACTCSNNIVDCRGKGTEIPTNLP-ETTEIRLEQNTIKV1PPGAESPYKKLRRIDL 335		16.9%; Score: 336; DB 4;	Pred No: 1 9e-15;							
Homo_sapiens (Human). Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. NCBI_TaxID: 9606; [1]	QY 109 QNNKIKEKENDFKGTLQFLGILNNNK1LK1HPRALFTKKLRLYLHSNQLSEPLNL 168		16.9%; Score: 336; DB 4;	Pred No: 1 9e-15;							
RN [2]	Db 336 SNNQISELAQDAFOGLRSLSLVSLYGNKTELPKS1LFEGLFSQ1LLINANK1NCLRVDIA 395		16.9%; Score: 336; DB 4;	Pred No: 1 9e-15;							
SEQUENCE FROM N.A. TISSUE=BRAIN; MEDLINE=99200389; PubMed=10102266; Wang K.H., Brose K., Arnott D., Kidd T., Goodman C.S., Henzel W., Tessier-Lavigne M.; "Purification of an axon elongation- and branch-promoting activity from brain identifies a mammalian Slit protein as a positive regulator of sensory axon growth.", Cell 96:771-784 (1999).	QY 169 PKSLAELR -- THENVKV1QKD1PKGMALHVLEMANS-----LDNNGK1T 212		16.9%; Score: 336; DB 4;	Pred No: 1 9e-15;							
RN [2]	Db 396 FQDLHNLLNLLSDYDNLQQTIAKGTESPLRA1QTMHLAQNPFICDCHLKRWLADYLTNPIE 455		16.9%; Score: 336; DB 4;	Pred No: 1 9e-15;							
SEQUENCE FROM N.A. TISSUE=BRAIN; MEDLINE=99200391; PubMed=10102268; Brose K., Blund K.S., Wang K.H., Arnott D., Henzel W., Goodman C.S., Tessier-Lavigne M., Kidd T.; "Slit proteins bind Robo receptors and have an evolutionary conserved role in repulsive axon guidance.", Cell 96:795-806 (1999).	QY 213 -PGA-----FEGYTVFH 223		16.9%; Score: 336; DB 4;	Pred No: 1 9e-15;							
RN [2]	Db 456 TSGARCSTSPRLANKRIGQ1KSKKFRCSGTEDYRSKLSDGCFADLACPEKCRCEGTV-- 513		16.9%; Score: 336; DB 4;	Pred No: 1 9e-15;							
RESULT 14 Q9Y5Q7 PRELIMINARY; PRT: 1525 AA.	QY 224 IRIABAKLTSVPKGKLDPTLHDYDYNK1STVELED-FRKYKELORLGLGNK1TIDJING 282		16.9%; Score: 336; DB 4;	Pred No: 1 9e-15;							
AC_09Y5Q7; DT_01-NOV-1999 (TREMBLrel. 12, Created) DT_01-NOV-1999 (TREMBLrel. 12, Last sequence update) DT_01-JUN-2002 (TREMBLrel. 21, Last annotation update) DE_SLT2.	Db 514 -DCSNQKLNKPHEHPOQYTAELRNINNFTYLEATG1KFQLPKR1KINFNSNKTIDIEG 572		16.9%; Score: 336; DB 4;	Pred No: 1 9e-15;							
Homo_sapiens (Human). Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. NCBI_TaxID: 9606; [1]	QY 283 SLANIPRVEREHLNNKLK -- KIPSG1PELKLQYQ1LFLHSN1STARVGSYNDCP1VKPM 338		16.9%; Score: 336; DB 4;	Pred No: 1 9e-15;							
RN [2]	Db 573 AFGEGSGVNELLTSR1LNQVQHMKFQGLESLKLTL--MLRSNRT1TCVGNDSF---- 623		16.9%; Score: 336; DB 4;	Pred No: 1 9e-15;							
SEQUENCE FROM N.A. TISSUE=BRAIN; MEDLINE=99200391; PubMed=10102268; Brose K., Blund K.S., Wang K.H., Arnott D., Henzel W., Goodman C.S., Tessier-Lavigne M., Kidd T.; "Slit proteins bind Robo receptors and have an evolutionary conserved role in repulsive axon guidance.", Cell 96:795-806 (1999).	QY 339 KKSLSYSAISLNFNPQKYWEMOPATERFCVLSRMSVQL 374		16.9%; Score: 336; DB 4;	Pred No: 1 9e-15;							
RN [2]	Db 624 GLSSVRLLSLYDNQQT--TVAPGAFDTLHSLSTLNL 657		16.9%; Score: 336; DB 4;	Pred No: 1 9e-15;							
RESULT 15 O94813 PRELIMINARY; PRT: 1529 AA.	DR 094813 PRELIMINARY; PRT: 1529 AA.		16.9%; Score: 336; DB 4;	Pred No: 1 9e-15;							
EMBL; AF33270; ADD25539.1; HSSP; P00743; JCCF	AC 094813; PRELIMINARY; PRT: 1529 AA.		16.9%; Score: 336; DB 4;	Pred No: 1 9e-15;							
DR InterPro; IPR00152; Asx_hydroxy_1.	AC 094813; PRELIMINARY; PRT: 1529 AA.		16.9%; Score: 336; DB 4;	Pred No: 1 9e-15;							
InterPro; IPR000359; Cys_knot.	DR 01-MAY-1999 (TREMBLrel. 10, Created)		16.9%; Score: 336; DB 4;	Pred No: 1 9e-15;							
DR InterPro; IPR000561; EGF_1-like.	DR 01-MAY-1999 (TREMBLrel. 10, Last sequence update)		16.9%; Score: 336; DB 4;	Pred No: 1 9e-15;							
DR InterPro; IPR000742; EGF_2.	DR 01-JUN-2002 (TREMBLrel. 21, Last annotation update)		16.9%; Score: 336; DB 4;	Pred No: 1 9e-15;							
DR InterPro; IPR001881; EGF_Ca.	DE SL1T-2 protein.		16.9%; Score: 336; DB 4;	Pred No: 1 9e-15;							
DR InterPro; IPR003645; FOLN.	GN Homo_sapiens (Human).		16.9%; Score: 336; DB 4;	Pred No: 1 9e-15;							
DR InterPro; IPR001791; Laminin_G.	OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. NCBI_TaxID: 9606; [1]		16.9%; Score: 336; DB 4;	Pred No: 1 9e-15;							
DR InterPro; IPR001611; LRR.	RP SEQUENCE FROM N.A. MEDLINE=903307; PMID=9813312; RA Itoh A., Miyabayashi T., Ohno M., Sakano S., Cloning and expression of three mammalian homologues of drosophila slit suggest possible roles for slit in the formation and maintenance of the nervous system.;		16.9%; Score: 336; DB 4;	Pred No: 1 9e-15;							
DR InterPro; IPR000483; LRR_Cterm.	RA Itoh A., Miyabayashi T., Ohno M., Sakano S., Cloning and expression of three mammalian homologues of drosophila slit suggest possible roles for slit in the formation and maintenance of the nervous system.;		16.9%; Score: 336; DB 4;	Pred No: 1 9e-15;							
DR InterPro; IPR000372; LRR_Nterm.	RA Itoh A., Miyabayashi T., Ohno M., Sakano S., Cloning and expression of three mammalian homologues of drosophila slit suggest possible roles for slit in the formation and maintenance of the nervous system.;		16.9%; Score: 336; DB 4;	Pred No: 1 9e-15;							
DR InterPro; IPR003592; LRR_out.	RA Itoh A., Miyabayashi T., Ohno M., Sakano S., Cloning and expression of three mammalian homologues of drosophila slit suggest possible roles for slit in the formation and maintenance of the nervous system.;		16.9%; Score: 336; DB 4;	Pred No: 1 9e-15;							
DR InterPro; IPR003591; LRR_typ.	RA Itoh A., Miyabayashi T., Ohno M., Sakano S., Cloning and expression of three mammalian homologues of drosophila slit suggest possible roles for slit in the formation and maintenance of the nervous system.;		16.9%; Score: 336; DB 4;	Pred No: 1 9e-15;							
DR InterPro; IPR0008; EGF_9.	RA Itoh A., Miyabayashi T., Ohno M., Sakano S., Cloning and expression of three mammalian homologues of drosophila slit suggest possible roles for slit in the formation and maintenance of the nervous system.;		16.9%; Score: 336; DB 4;	Pred No: 1 9e-15;							
DR InterPro; IPR0054; Laminin_G; 1.	RA Itoh A., Miyabayashi T., Ohno M., Sakano S., Cloning and expression of three mammalian homologues of drosophila slit suggest possible roles for slit in the formation and maintenance of the nervous system.;		16.9%; Score: 336; DB 4;	Pred No: 1 9e-15;							
DR InterPro; IPR00560; LRR; 17.	RA Itoh A., Miyabayashi T., Ohno M., Sakano S., Cloning and expression of three mammalian homologues of drosophila slit suggest possible roles for slit in the formation and maintenance of the nervous system.;		16.9%; Score: 336; DB 4;	Pred No: 1 9e-15;							
DR InterPro; IPR0463; LRRCT; 4.	RA Itoh A., Miyabayashi T., Ohno M., Sakano S., Cloning and expression of three mammalian homologues of drosophila slit suggest possible roles for slit in the formation and maintenance of the nervous system.;		16.9%; Score: 336; DB 4;	Pred No: 1 9e-15;							
DR InterPro; IPR0462; LRRCT; 4.	RA Itoh A., Miyabayashi T., Ohno M., Sakano S., Cloning and expression of three mammalian homologues of drosophila slit suggest possible roles for slit in the formation and maintenance of the nervous system.;		16.9%; Score: 336; DB 4;	Pred No: 1 9e-15;							
DR SMART; SM00041.	RA Itoh A., Miyabayashi T., Ohno M., Sakano S., Cloning and expression of three mammalian homologues of drosophila slit suggest possible roles for slit in the formation and maintenance of the nervous system.;		16.9%; Score: 336; DB 4;	Pred No: 1 9e-15;							
DR SMART; SM00179; EGF_CA; 2.	RA Itoh A., Miyabayashi T., Ohno M., Sakano S., Cloning and expression of three mammalian homologues of drosophila slit suggest possible roles for slit in the formation and maintenance of the nervous system.;		16.9%; Score: 336; DB 4;	Pred No: 1 9e-15;							

DR	InterPro: IPR000359; CYS_knot.
DR	InterPro: IPR000561; EGF-like.
DR	InterPro: IPR000742; EGF_2.
DR	InterPro: IPR000881; EGF_Ca.
DR	InterPro: IPR000645; FOLN.
DR	InterPro: IPR000791; Laminin_G.
DR	InterPro: IPR000611; LRR.
DR	InterPro: IPR000483; LRR_Cterm.
DR	InterPro: IPR000372; LRR_Nterm.
DR	InterPro: IPR000592; LRR_out.
DR	InterPro: IPR000591; LRR_TYP.
DR	Pfam: PF00008; BCF; 9.
Pfam	PF00054; laminin_G; 1.
DR	Pfam: PF00560; LRR; 17.
Pfam	PF01463; LRRCT; 4.
Pfam	PF01462; LRRNT; 4.
SMART	SM00041; CT; 1.
DR	SMART: SM00079; EGF_Ca; 2.
SMART	SM00001; EGF_like; 7.
DR	SMART: SM00274; FOLN; 2.
SMART	SM00282; LamG; 1.
DR	SMART: SM00370; LRR; 6.
SMART	SM00082; LRRCT; 4.
SMART	SM00369; LRRN; 4.
DR	PROSITE: PS00010; ASX_HYDROXYL; UNKNOWN_2.
DR	PROSITE: PS01187; CNCK_1; UNKNOWN_1.
DR	PROSITE: PS01225; CNCK_2; 1.
DR	PROSITE: PS00022; EGF_1; UNKNOWN_9.
DR	PROSITE: PS01187; EGF_2; 7.
DR	PROSITE: PS01187; EGF_Ca; 2.
DR	Calcium-binding, EGF-like domain; Glycoprotein; Repeat.
ISQ	SEQUENCE 1529 AA: 169869 MW;
ISQ	SEQUENCE 1529 AA: 5D19CC5E7FD461BA CRC64;
Query Match	Length 15
Best Local Similarity	Score 336; DB 4;
Matches 105	pred. No. 1.9e-15;
Conservative	N mismatches 120; Indels
74	CPEGVYQSRVHVCSDSLGLUTSPVNTPFDT-
273	CPAACTCSNNVDCRGKGLTEIPNLP-ETITEIRLEONTTIVKIPPOAFSPVY
Qy	109 QNNKTEIKENDEFKGLTSLYGLILNNNNKLTKHPKAFTTKLRLRLLYLSHNC
Ddb	332 SNNQISELAPIAFOGLRSNSLNLVLYGNKNTTEPLKSFLFEGLFSQQLLNLNANK
Qy	169 PKS1AFLR--IHENKVKKIQQDKTFKGMMALHYLEMSANP--
Ddb	392 FQDLHNLLNLSLYDNKLQTAAKGTSPLRAIQTMHLAQQNPFICDCH1KWLAD
213	--
452	TSGARTSPRLLANKRIGQIKSKKFRCSAKEQYFIPGTEDYSKLSGDCFAD
Qy	216 FEGTVYFHRIEAEEKLTSVPKGKLPPTLLEHLDYNKISTVELED-FKRYKEL
Ddb	512 CRGTVW--DGSNQNLK1PEHTQYTAELRNNEPTVLEATGFKKLPOOL
Qy	275 KTDIENGSLNANTPVREHLENNKL--KPSGGLPELKVLQI1FLHSNS
Ddb	569 KTDIIEGAAFGASGSYNEITLTSRNLVYQHMKGLESLKLTL--MLRSNR
214	--
452	TSGARTSPRLLANKRIGQIKSKKFRCSAKEQYFIPGTEDYSKLSGDCFAD
Qy	216 FEGTVYFHRIEAEEKLTSVPKGKLPPTLLEHLDYNKISTVELED-FKRYKEL
Ddb	512 CRGTVW--DGSNQNLK1PEHTQYTAELRNNEPTVLEATGFKKLPOOL
Qy	275 KTDIENGSLNANTPVREHLENNKL--KPSGGLPELKVLQI1FLHSNS
Ddb	569 KTDIIEGAAFGASGSYNEITLTSRNLVYQHMKGLESLKLTL--MLRSNR
214	--
331	FCPTVPKMKSSLYSAISLFNPVKWMPATERCVLRSMSYVQ 374
F	--IGLSVRLSLSDNQIT-TVAAGAFTDLISLSTNL 661
626	F--

Result No.	Score	Query Match	Length	DB ID	Description
1	104.5	52.2	369	S32793	biglycan precursor - m proteoglycan 1 - m biglycan precursor
2	103.5	52.1	369	S32793	biglycan precursor
3	103.4	51.9	368	BGIU1	biglycan precursor
4	102.5	51.4	368	S32559	biglycan precursor
5	96.9	48.6	357	S24317	decorin precursor
6	96.3	48.4	359	NHUUC8	decorin precursor
7	94.3	47.3	360	S06280	decorin precursor
8	94.1	47.2	360	147020	decorin - rabbit
9	907.5	45.6	354	A55454	decorin precursor
10	89.9	45.1	354	S29145	decorin precursor
11	399.5	20.1	343	A41748	lumican precursor
12	384	19.3	342	A46743	lumican precursor
13	378	19.0	338	S52284	lumican, secretory protein, arginine-fibromodulin precursor
14	358.5	18.0	382	139068	lumican-like protein 2 precursor
15	342.5	17.2	375	S05390	fibromodulin precursor
16	333.5	16.7	380	S71876	fibromodulin - chi MEGF5 protein - raf
17	324	16.3	1523	T13953	slit-1 protein homolog
18	323	16.2	1531	T42218	fibromodulin precursor
19	318	16.0	376	S55275	slit protein 2 precursor
20	316	15.9	1469	B36665	slit protein 1 precursor
21	316	15.9	1480	A36665	slit protein 1 precursor
22	296	14.9	1091	A58532	gial cell membrane chondroderin precursor
23	294	14.8	361	A53862	secreted leucine-rich insulin-like growth factor-binding protein
24	282.5	14.2	1025	T42626	neuronal leucine-rich insulin-like growth factor-binding protein
25	272	13.7	603	JC1282	insulin-like growth factor-binding protein
26	266	13.4	707	JC7763	insulin-like growth factor-binding protein
27	265	13.3	603	JC6128	insulin-like growth factor-binding protein
28	265	13.3	1051	T13174	insulin-like growth factor-binding protein
29	261.5	13.1	907	JO1913	insulin-like growth factor-binding protein

Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2003, 16:01:11 ; Search time 22 Seconds (without alignments)

Title: US-09-944-457-2

Perfect score: 1992

Sequence: 1 MKEYVLLFLAICSAKPFES.....PATFRCVLSRMSVOLGNFGM 379

Scoring table: BLOSUM62

Gapopen: 10.0 , Gapext: 0.5

Searches: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR\_73;\*

1: pir1;\*

2: pir2;\*

3: pir3;\*

4: pir4;\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Qry	Match	Length	DB ID	Score	Best Local Similarity	DB 2:	Length	Score	Best Local Similarity	DB 2:	Length
1	52.2	369	S32793	52.28	FSPSPHIAKLNMLKDMEDDDDDDDDDDDDDNSLFPTR	52.28	369;	1040.5	52.28	52.28	369;
2	52.1	369	S32793	52.28	WLTLLIASQLPFEOKFWDETLDDGLLMNDEEASGSDTTSGVPDLID	52.28	63;	1.2e-63;	52.28	52.28	63;
3	51.9	368	BGIU1	52.28	Mishatches 67; Indels 19; Gaps 3;	52.28	63;	1.2e-63;	52.28	52.28	63;
4	51.4	368	S32559	52.28	YVILLFLALCSAKPF-----FSPSPHIAKLNMLKDMEDDDDDDDDDDDNSLFPTR	52.28	63;	1.2e-63;	52.28	52.28	63;
5	48.6	357	S24317	52.28	WLTLLIASQLPFEOKFWDETLDDGLLMNDEEASGSDTTSGVPDLID	52.28	63;	1.2e-63;	52.28	52.28	63;
6	48.4	359	NHUUC8	52.28	EPRSHEFFPEPDLEPMCPGQCYSRVVHCSIDGLTSVPTNPFDTRMLDQNKKIREKTEK 119	52.28	63;	1.2e-63;	52.28	52.28	63;
7	47.3	360	S06280	52.28	60	52.28	63;	1.2e-63;	52.28	52.28	63;
8	47.2	360	147020	52.28	-----FSAMCPGCHCLRVQQCSDLGLKTVPEKISPDPTLLDQNNDISLRKD 109	52.28	63;	1.2e-63;	52.28	52.28	63;
9	45.6	354	A55454	52.28	Qy 4 YVILLFLALCSAKPF-----FSPSPHIAKLNMLKDMEDDDDDDDDDDDNSLFPTR	52.28	63;	1.2e-63;	52.28	52.28	63;
10	45.1	354	S29145	52.28	Db 5 WLTLLIASQLPFEOKFWDETLDDGLLMNDEEASGSDTTSGVPDLID	52.28	63;	1.2e-63;	52.28	52.28	63;
11	399.5	20.1	A41748	52.28	Qy 5 WLTLLIASQLPFEOKFWDETLDDGLLMNDEEASGSDTTSGVPDLID	52.28	63;	1.2e-63;	52.28	52.28	63;
12	384	19.3	A46743	52.28	Db 60 EPRSHEFFPEPDLEPMCPGQCYSRVVHCSIDGLTSVPTNPFDTRMLDQNKKIREKTEK 119	52.28	63;	1.2e-63;	52.28	52.28	63;
13	378	19.0	S52284	52.28	Db 60 -----FSAMCPGCHCLRVQQCSDLGLKTVPEKISPDPTLLDQNNDISLRKD 109	52.28	63;	1.2e-63;	52.28	52.28	63;
14	358.5	18.0	139068	52.28	Qy 60 DFKGLTSLIGYLILNNKLTHPKAFITTKRLYLSHNLSEPLNLKSLAELRIRHE 179	52.28	63;	1.2e-63;	52.28	52.28	63;
15	342.5	17.2	S05390	52.28	Db 110 DFKGLQHLYALVNNKISKIHEKAFLSPRLKLQKLYSKNHLVEPPNLSLVELRIRD 169	52.28	63;	1.2e-63;	52.28	52.28	63;

Total number of hits satisfying chosen parameters:

283224

## ALIGNMENTS

Qy	180 NKKKIQDKDFFKGMMNALHVLEMSANPLDNGIIEPGAFEGTVPHIRIAEAKLTSVPKGLP 239	Oy	4 YVLLFLALCSAKPF ---FSPSHIALKNNMLKDMDTDDDDDEDNSLFPT 59
Db	170 NTRKVKPGYFGSLRNMCNTEGMQGK 229	Db	5 WLTLLJALSGALPFFQKGWDFDTLDDGLMMNDEASGGDTTSGPVDLO ---SVPPT 59
Qy	240 PTLLELHDYDKNSTVLEDKRYKELQRGLGNKITTDLENGSLANIPVRTHLENNK 299	Qy	60 EPRSHFFPDIDLPMPMCQPGCQCSRVVHCSIDLGLTSPVTPNLPFDTRMLDLQNNNKEIKEKEN 119
Db	230 ETLNELHLHDNKTQATELDLRYSKLYRLGLGQINQIRGMGSLSFLPTLRLHLDNPK 289	Db	60 -----FSAMCPGCGCHLRVQCSIDLGLKVPKEISPDPTLQLDQNNDISLRKD 109
Qy	300 LKKIPSGLPEIJKLQILFLHSNSTARVQWNDPFCPTVPMKKSQIYSAISLFNNVQYHMEMO 359	Qy	120 DFKGLTSLYGLUILLNNKKLTKIHPKAFLTTKLRRLYLSHQNSEIPLNPKSLAERTRIE 179
Db	290 LSRVPAGLPDKLQVYLVNNKISKHEKATSPRLQKLYLSKKNHVEPPNLSSVLVELRHD 169	Db	110 DFKGLQHAYALVNNKISKHEKATSPRLQKLYLSKKNHVEPPNLSSVLVELRHD 169
Qy	360 PATFRCVLSRMSVQLGNF 377	Qy	180 NKVKKIQDKDTEKGMMNALHVLEMSANPLDNGIIEPGAFEGTVPHIRIAEAKLTSVPKGLP 239
Db	350 PATFRCVTDRLAIQFGNY 367	Db	170 NRTRKVQKGVEFGGLRNINCIEGNGPNLENSQEPGAFTGKLNRYURISEAKLTGIPKDLP 229
<hr/>			
RESULT 2			
S20811	proteoglycan I - mouse	Qy	240 PTLLELHDYDKNSTVLEDKRYKELQRGLGNKITDENGSLANIPVRVRETHLENNK 299
N: Alternative names: biglycan		Db	230 ETLNELHLHDNKTQATELDLRYSKLYRLGLGHNQTRMENGSLSPFLPTLRLHLDNPK 289
R: Species: Mus musculus (house mouse)	Qy	300 LKKIPSGLPEIJKLQILFLHSNSTARVQWNDPFCPTVPMKKSQIYSAISLFNNVQYHMEMO 359	
Date: 20-Feb-1995 #sequence_change 20-Aug-1999	Db	290 LSRVPAGLPDKLQVYLVNNKISKHEKATSPRLQKLYLSKKNHVEPPNLSSVLVELRHD 169	
Accession: S20811; A57645; 149534	Qy	360 PATFRCVLSRMSVQLGNF 377	
R: Naitoh, Y.; Suzuki, S.	Db	350 PATFRCVTDRLAIQFGNY 367	
Submitted to the EMBL Data Library, July 1990			
A: Description: Nucleotide sequences of cDNAs encoding mouse PGII and PGIII.			
A: Reference number: S20811			
A: Status: preliminary			
A: Molecule type: mRNA			
A: Residues: 1-369 <NAI>			
A: Cross-references: EMBL:X53928; PIDN:CAA37875.1; PID:953667			
R: Negrowski, Y.; Pillarisetti, J.; Danielson, K.G.; Suzuki, S.; Iozzo, R.V.			
Genomics 30, 8-17, 1995			
A: Title: The murine biglycan: complete cDNA cloning, genomic organization, promoter function			
A: Reference number: A57645; MUID:96129295; PMID:8555907			
A: Accession: A57645			
A: Status: preliminary			
A: Molecule type: mRNA			
A: Residues: 1-67 'W, 69-369 <WEC>			
A: Cross-references: GB:L20276; PIDN:AAA64360.1; PID:9348962			
R: Rau, W.; Just, W.; Vetter, U.; Vogel, W.			
Mamm. Genome 5, 395-396, 1994			
A: Title: A dinucleotide repeat in the mouse biglycan gene (Bst) on the X chromosome.			
A: Reference number: 149534; MUID:94319093; PMID:8043960			
A: Accession: 149534			
A: Status: preliminary			
A: Molecule type: mRNA			
A: Residues: 1-67 'W, 69-369 <RES>			
A: Cross-references: GB:L20276; PIDN:9348961; PID:9348962			
C: Genetics:			
A: Gene: Bgn			
C: Superfamily: decorin; leucine-rich alpha-2-glycoprotein repeat homology; proteoglycan			
C: Keywords: chondroitin sulfate proteoglycan; dermatan sulfate; extracellular matrix; glycoprotein			
F: 58-82/Domain: proteoglycan amino-terminal homology <PH>			
F: 116-119/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR1>			
F: 140-160/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR2>			
F: 185-208/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR3>			
F: 210-230/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR4>			
F: 231-255/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR5>			
F: 255-278/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR6>			
F: 279-301/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR7>			
F: 302-316/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR8>			
F: 317-369/Domain: proteoglycan carboxy-terminal homology <LR9>			
Query Match 52.1%			
Best Local Similarity 52.4%			
Matches 198; Conservative 68; Mismatches 93; Indels 19; Gaps 3;			
A: Experimental source: aorta			
A: Residues: 38-57 <S10>			
A: Experimental source: aorta			

Roughley, P.J.; White, R.J.  
Biochem. J. 262; 823-827, 1989  
A; Reference number: S05639; MUID:90073579; PMID:2590169  
A; Molecule type: protein  
A; Accession: S05639  
A; Residues: 38-41, 'X', 43-46, 'X', 48-57 <ROU>  
R; Fisher, L.W.; Hawkins, G.R.; Tuross, N.; Termine, J.D.  
J. Biol. Chem. 262, 9702-9708, 1987  
A; Title: Purification and partial characterization of small proteoglycans I and II, bone  
A; Reference number: A92656; MUID:87250639; PMID:3597437  
A; Accession: A28457  
A; Molecule type: protein  
A; Residues: 38-41, 'X', 43-62, 'X', 64-66 <FIGS>  
A; Experimental source: bone  
C; Genes:  
A; Gene: GDB:BGN  
A; Cross-references: GDB:119727; OMIM: 301870  
A; Map position: Xq8-Xq8  
A; Introns: 80/1; 117/3; 189/1; 226/1; 257/2; 303/3  
C; Superfamily: decorin; leucine-rich alpha-2-glycoprotein repeat homology; proteoglycan  
F; 1-16/Domain: signal sequence #status predicted <STG>  
F; 17-37/Domain: chondroitin sulfate proteoglycan; dermatan sulfate; duplication; extracellular  
F; 57-81/Domain: leucine-rich alpha-2-glycoprotein repeat homology <PAH>  
F; 91-114/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>  
F; 115-138/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>  
F; 139-159/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>  
F; 160-183/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>  
F; 184-207/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>  
F; 209-229/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>  
F; 220-253/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR7>  
F; 254-277/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR8>  
F; 278-300/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR9>  
F; 301-315/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR10>  
F; 316-368/Domain: proteoglycan carboxy-terminal homology #status atypical <LR>  
F; 42-47/Binding site: dermatan sulfate (Ser) (covalent) #status experimental  
F; 180, 198/Binding site: dermatan sulfate (Asn) (covalent) #status predicted  
F; 270, 311/Binding site: carbohydrate (Asn) (covalent) #status predicted  
Query Match 51.9% Score 1034; DB 1; Length 368;  
Best Local Similarity 53.2%; Pred. No. 3.4e-63; Matches 200; Conservative 63; Mismatches 93; Indels 20; Gaps 3;  
Matches 200; Conservatve 63; Mismatches 93; Indels 20; Gaps 3;

6 LILFLALCSAKP-----FSPSHLALKNNMILKMDMDDDDDDDDDDDDDDDDDDDDNLFPTREP 61  
7 LYSDLAISQALFEQRGFWDFLDDGPFFMMNDEEAGADTSVQLDPD-----SVPTPYs- 60

Qy 62 RSHFFPDLFPCPFGQCYSEVSFHSDGLTSVPNTNIPFDRLQNNRKTEIKENDF 121  
Do 61 -----ANCPFGCHLFRVQVSDIGLKSVKEITSDFLDLQNNKTSERDKDDF 110

Qy 122 KGLTSLYGLUNNKLTQTKLPRPLFTKKLRLRPLTGHNLQNSEIPLNPKSLAELRHENK 181  
Db 111 KGQHLYALVNKNKSKIHERAFSPSLRKQKLYLSKHNHIVEIPNLPSSLEVRIHNDR 170

Qy 182 VKKTOKDTEKGNAHVLEMNSANPLDNNGLEPAGFEGTVYHTRIAEAKITSVKGKGLPPT 241  
Db 171 IKRVPKCVESGGRNMCIEMGGNPLENSGEFGAPGDLKLNYLRLSEAKLTGFKDLPET 230

Oy 242 LFLFLHDYNKISTVELEDFKRYKELORLIGLNKTFDIENGSLANTPVREHLENKLIK 301  
Db 231 LNLBLHDHNKCAIEFLDLJRYSKLYRQKLYGHNQFMRMIEGNSLSPLETPRELHDNNKLA 290

Qy 302 KIPSGSPPELYKLQYLIFHSNTRAVGNDPCPTVPMKMKSLYSALSLENNPVKYWEMOPA 361  
Do 291 RPSGLPDLKLQVYVLLHSNNITKVGNDFCPMGFGVKRAYINGISLFFFNPVPPWEVQPA 350

Oy 362 TFRCLVSRMSVQGKF 377  
Db 351 TFRCTVTDRLAQFGNY 366

RESULT 4  
S32559  
biglycan precursor - bovine  
N; Alternative names: dermatan sulfate proteoglycan I (DS-PGI); proteochondroitin core p  
C; Species: Bos primigenius taurus (cattle)  
C; Date: 03-May-1994 #sequence revision 20-Feb-1995 #text change 07-May-1999  
C; Accession: S32559; S34229; A33701; A31430; PT0078; S55673; A33137  
R; Tork, M.A.; Evans, S.A.S.; Marcum, J.A.  
Biochim. Biophys. Acta 1173, 81-84, 1993  
A; Title: cDNA sequence for bovine biglycan (PGI) protein core.  
A; Reference number: S32559; MUID:9325052; PMID:8485158  
A; Accession: S32559  
A; Molecule type: mRNA  
A; Residues: 1-369 <PRO>  
A; Cross-references: EMBL:1.07953; MIBI:9162746  
A; Experimental source: aortic smooth muscle  
R; Marcum, J.A.; Tork, M.; Evans, S.  
Submitted to the EMBL Data Library, December 1992  
A; Reference number: S34229  
A; Accession: S34229  
A; Molecule type: mRNA  
A; Residues: 1-250, 'V', 252-369 <MAR>  
A; Cross-references: EMBL:1.07953  
R; Neame, P.J.; Choi, H.U.; Rosenberg, L.C.  
J. Biol. Chem. 264, 8653-8661, 1989  
A; Title: the primary structure of the core protein of the small, leucine-rich proteoglycan  
A; Reference number: A33701; MUID:89255324; PMID:2656687  
A; Molecule type: protein  
A; Accession: A33701  
A; Molecule type: protein  
A; Residues: 36-187, 'E', 189-367, 'Y' <NEA>  
A; Experimental source: cartilage  
R; Choi, H.U.; Johnson, T.L.; Pal, S.; Tang, L.H.; Neame, P.J.  
J. Biol. Chem. 264, 2876-2884, 1989  
A; Title: Characterization of the dermatan sulfate proteoglycans, DS-PGI and DS-PGII,  
A; Reference number: A31430; MUID:89123388; PMID:2914936  
A; Accession: A31430  
A; Molecule type: protein  
A; Residues: 38-41, 'X', 43-47, 'X', 49-63 <CHO>  
A; Note: sequences from skin and cartilage were identical  
R; Marcum, J.A.; Thompson, M.A.  
Biochem. Biophys. Res. Commun. 175, 706-712, 1991  
A; Title: The amino-terminal region of a proteochondroitin core protein, secreted by a  
om human bone.  
A; Reference number: PT0078; MUID:91207372; PMID:2018513  
A; Accession: PT0078  
A; Molecule type: protein  
A; Residues: 17-24, 'F', 26-30 <MA2>  
A; Experimental source: aortic smooth muscle  
R; Scott, P.G.; Nakano, T.; Dodd, C.M.  
Biochim. Biophys. Acta 1244, 121-128, 1995  
A; Title: Small proteoglycans from different regions of the fibrocartilaginous temporo  
A; Reference number: S55673; MUID:95284673; PMID:7766647  
A; Accession: S55673  
A; Molecule type: protein  
A; Residues: 38-41, 'X', 43-47, 'X', 49-53 <SCD>  
C; Super family: decorin; leucine-rich alpha-2-glycoprotein repeat homology; proteoglycan  
F; 1-16/Domain: signal sequence #status predicted <SI>  
F; 17-37/Domain: amino-terminal peptide; biglycan #status predicted <MAT>  
F; 38-369/Domain: proteoglycan amino-terminal homology <PAH>  
F; 58-82/Domain: proteoglycan repeat homology <LRR1>  
F; 92-115/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>  
F; 115-139/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>  
F; 140-160/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>  
F; 161-184/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>  
F; 185-208/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>  
F; 210-230/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>  
F; 231-254/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR7>  
F; 255-278/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR8>  
F; 279-301/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR9>  
F; 302-316/Domain: leucine-rich alpha-2-glycoprotein repeat homology #status atypical  
F; 317-365/Domain: proteoglycan carboxyl-terminal homology <PC1>



Biochem. J. 262, 823-827, 1989  
 A;Title: Dermatan sulphate proteoglycans of human articular cartilage. The properties of  
 A;Reference number: S05639; MUID:9007579; PMID:2590169  
 A;Accession: S05640  
 A;Molecule type: protein  
 A;Residues: 31-33,'X',35-50 <ROU>  
 C;Comment: This protein binds type I collagen.  
 C;Genetics:  
 A;Gene: GDB:DCN  
 A;Cross-references: GDB:119939; OMIM:125255  
 A;Map position: 1q21.3-1q23  
 A;Introns: 71/1; 108/3; 180/1; 218/1; 249/2; 295/3  
 A;Note: the first two introns occur before the initiator codon  
 C;Superfamily: decorin; leucine-rich alpha-2-glycoprotein repeat homology; proteoglycan  
 F;1-16/Domain: chondroitin sulfate proteoglycan; collagen binding; dermatan sulfate; dupli-  
 E;17-30/Domain: propeptide #status predicted <SIG>  
 E;31-359/Product: decorin #status predicted <PRO>  
 F;38-72/Domain: proteoglycan amino-terminal homology <PAH>  
 F;78-105/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>  
 F;106-129/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>  
 F;130-150/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>  
 F;151-174/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>  
 F;175-200/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>  
 F;201-221/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>  
 F;222-245/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR7>  
 F;246-269/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR8>  
 F;270-292/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR9>  
 F;293-316/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR10>  
 F;34/Binding site: dermatan sulfate (Ser) (covalent) #status experimental  
 F;189,355/Binding site: carbohydrate (Asn) (covalent) #status predicted  
 F;211,262,303/Binding site: carbohydrate (Ser) (covalent) #status predicted

Query Match 48.4% Score 963.5; DB 1; Length 359;  
 Best Local Similarity 50.4% Pred. No. 2,1e-58;  
 Matches 191; Conservative 65; Mismatches 100; Indels 23; Gaps 4;

Qy 1 MKEVYLLPLALCS-AKPEFFSPSHIALKNMMLKDMEDTDDDDDDDDDDDDNSELFPTR 59  
 Db - 1 MKATILLILAQSVWAGPF -- QQRGLFDFMIDEASGIGPVEPVDFD----- 45  
 Qy 60 EPRSHFFPFDFLPMCPGCOCTSYRVYHCSLDLGLTSYPTNIPFDTRMLDLQNNKIKEIKEN 119  
 Db 46 -----FEPSSLGPGVCPFRQCQHLVRVQCSIDLQKVDLPPDTTLIDLQNNKITEIKDG 99  
 Qy 120 DFGKLTSYGLLNNNNKLTKIHPKAFTLTTKLRRLYLHSNOSETPLNPKSLAERIHE 179  
 Db 100 DFLNLKNEHALVLVNNKISKVSPGAFTPLVKLERYLTSKQNLPEKMKLTQLRAHE 159  
 Qy 180 NIVKVKIOKDTEFGMMALHVRMLVHNSANPLDNGTEPQAFEGV-TVFHRIIAFKLTSPVPKGL 238  
 Db 160 NEITKVKVTFTGFLNQMVIEGTNPQKSGGTENGAFQGMKRLSYRIATNTITSIPQGL 219  
 Qy 239 PPTLLELHDLYNKISTWELEDFKRYKELQRLGLGNKNTIDIEGLSLANTPRVRETHLENN 298  
 Db 220 PPSLTELHLDGKRSIRDAASLKGLNNLAKLGSFSNSISAVDGLSLANTPHRELHLDNN 279  
 Qy 299 KUKKIPGGLPEKYLQITFLHSNSIARVGYNDFCPIPVPKMRKSLYSIAISLNFNPVQYWEI 358  
 Db 280 KLTTRVPGLAEEKYIQVYLVHNINNSVVGSSDFCPGHNTKKASYSGVSLNSNPVQYWEI 339  
 Qy 359 QPATFRGCVLSRMSVOLQNF 377  
 Db 340 QPSTFRCVYVRSMSVOLQNF 358

Query Match 47.3%; Score 943.; DB 2; Length 360;  
 Best Local Similarity 49.1%; Pred. No. 5.3e-57;  
 Matches 186; Conservative 69; Mismatches 102; Indels 22; Gaps 5;

Qy 1 MKEYVLLPLALCS-AKPEFFSPSHIALKNMMLKDMEDTDDDDDDDDNSELFPTR 59  
 Db 1 MKATILLILAQSVWAGPF -- QQRGLFDFMEL----- 48

RESULT 7  
 S06280  
 decorin precursor - bovine  
 N Alternative names: dermatan sulfate proteoglycan II; proteoglycan core protein II  
 C;Species: Bos primigenius taurus (cattle)  
 C;Date: 31-Mar-1990 #sequence\_revision 31-Mar-1990 #text\_change 20-Aug-1999

Qy 60 EPRSHFFPFDFLPMCPGCOCTSYRVYHCSLDLGLTSYPTNIPFDTRMLDLQNNKIKEIKEN 119  
 Db 49 Ep -----MGPVCPERQOCHRVQCSIDLQKVDLPPDTTLIDLQNNKITEKDG 100  
 Qy 120 DFKGILTSLYGLLNNNNKLTKIHPKAFTLTTKLRRLYLHSNOSETPLNPKSLAERIHE 179



Db	267	REHLDDNNKLLRVPAGLAQHKYIQVYVLHNNNNTSAYQNDFFCRAUGHPSRKASYSAVSLYG	326	Db	211	PQGLPITSSELHDGNKIAKYDAASLKGMSNLSKUGLSFNSITVVENGSLANVPHLRELH	270
Oy	351	NPVRYWEMOPATFRCVLSRMSJOLGNE	377	Oy	295	LENNKIKTPGGLPEPKYLOLIFHNSTARVGVDNFCPTVKUKKTSALISLFNNPK	354
Db	327	NPVRYWEIPNPFRCVTVRSAIOLGNY	353	Db	271	LDNNKKLRLRPAGLAQHKYQVYVLHNNNNTSEVGQHDFCLPSYQTRKTSTAVSLISNPR	330
RESULT 10				Qy	355	YWEMOPATFRCVLSRMSVOLGNF	377
S29145				Db	331	YQJHPTFRCVFGRTSIOLGNY	353
decorin precursor - rat				Db	211	PQGLPITSSELHDGNKIAKYDAASLKGMSNLSKUGLSFNSITVVENGSLANVPHLRELH	270
C;Species: Rattus norvegicus (Norway rat)				Oy	295	LENNKIKTPGGLPEPKYLOLIFHNSTARVGVDNFCPTVKUKKTSALISLFNNPK	354
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 20-Aug-1999				Db	271	LDNNKKLRLRPAGLAQHKYQVYVLHNNNNTSEVGQHDFCLPSYQTRKTSTAVSLISNPR	330
C;Accession: S29145; 160238; S29517				Qy	355	YWEMOPATFRCVLSRMSVOLGNF	377
R;Abramson, S.R.; Woessner Jr., J.F.				Db	331	YQJHPTFRCVFGRTSIOLGNY	353
Biochim. Biophys. Acta 1132, 225-227, 1992.				Db	211	PQGLPITSSELHDGNKIAKYDAASLKGMSNLSKUGLSFNSITVVENGSLANVPHLRELH	270
M;Sequence number: S29145; MUID:93003331; PMID:1390895				Oy	295	LENNKIKTPGGLPEPKYLOLIFHNSTARVGVDNFCPTVKUKKTSALISLFNNPK	354
Accession: S29145				Db	271	LDNNKKLRLRPAGLAQHKYQVYVLHNNNNTSEVGQHDFCLPSYQTRKTSTAVSLISNPR	330
A;Species: preliminary				Qy	355	YWEMOPATFRCVLSRMSVOLGNF	377
A;Residues: 1-354 <ABR>				Db	331	YQJHPTFRCVFGRTSIOLGNY	353
A;Cross-references: EMBL:212298; PID:957549; PIDN:CAA78170_1; PID:957550				Db	211	PQGLPITSSELHDGNKIAKYDAASLKGMSNLSKUGLSFNSITVVENGSLANVPHLRELH	270
R;Asundi, V.K.; Dreher, K.L.				Oy	295	LENNKIKTPGGLPEPKYLOLIFHNSTARVGVDNFCPTVKUKKTSALISLFNNPK	354
Eur. J. Cell Biol. 59, 314-321, 1992.				Db	271	LDNNKKLRLRPAGLAQHKYQVYVLHNNNNTSEVGQHDFCLPSYQTRKTSTAVSLISNPR	330
A;Title: Molecular characterization of vascular smooth muscle decorin: deduced core protein				Qy	355	YWEMOPATFRCVLSRMSVOLGNF	377
A;Reference number: 160238; MUID:93154359; PMID:1493796				Db	331	YQJHPTFRCVFGRTSIOLGNY	353
A;Accession: 160238				Db	211	PQGLPITSSELHDGNKIAKYDAASLKGMSNLSKUGLSFNSITVVENGSLANVPHLRELH	270
A;Status: preliminary				Oy	295	LENNKIKTPGGLPEPKYLOLIFHNSTARVGVDNFCPTVKUKKTSALISLFNNPK	354
A;Molecule type: mRNA				Db	271	LDNNKKLRLRPAGLAQHKYQVYVLHNNNNTSEVGQHDFCLPSYQTRKTSTAVSLISNPR	330
A;Residues: 11-354 <RES>				Qy	355	YWEMOPATFRCVLSRMSVOLGNF	377
A;Cross-references: EMBL:X59859; PID:956056; PIDN:CAA42519_1; PID:956057				Db	331	YQJHPTFRCVFGRTSIOLGNY	353
C;Genetics:				Db	211	PQGLPITSSELHDGNKIAKYDAASLKGMSNLSKUGLSFNSITVVENGSLANVPHLRELH	270
A;Gene: DCN				Oy	295	LENNKIKTPGGLPEPKYLOLIFHNSTARVGVDNFCPTVKUKKTSALISLFNNPK	354
C;Superfamily: decorin; leucine-rich alpha-2-glycoprotein repeat homology; proteoglycan				Db	271	LDNNKKLRLRPAGLAQHKYQVYVLHNNNNTSEVGQHDFCLPSYQTRKTSTAVSLISNPR	330
C;Keywords: collagen binding; extracellular matrix; glycoprotein				Qy	355	YWEMOPATFRCVLSRMSVOLGNF	377
F;31-354/Product: decorin #status predicted <SIG>				Db	331	YQJHPTFRCVFGRTSIOLGNY	353
F;43-67/Domain: proteoglycan amino-terminal homology <PAH>				Db	211	PQGLPITSSELHDGNKIAKYDAASLKGMSNLSKUGLSFNSITVVENGSLANVPHLRELH	270
F;70-101-124/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR1>				Oy	295	LENNKIKTPGGLPEPKYLOLIFHNSTARVGVDNFCPTVKUKKTSALISLFNNPK	354
F;125-145/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR2>				Db	271	LDNNKKLRLRPAGLAQHKYQVYVLHNNNNTSEVGQHDFCLPSYQTRKTSTAVSLISNPR	330
F;170-193/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR4>				Qy	355	YWEMOPATFRCVLSRMSVOLGNF	377
F;196-216/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR6>				Db	331	YQJHPTFRCVFGRTSIOLGNY	353
F;265-287/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR8>				Db	211	PQGLPITSSELHDGNKIAKYDAASLKGMSNLSKUGLSFNSITVVENGSLANVPHLRELH	270
F;303-354/Domain: proteoglycan carboxyl-terminal homology <PC1>				Oy	295	LENNKIKTPGGLPEPKYLOLIFHNSTARVGVDNFCPTVKUKKTSALISLFNNPK	354
Query Match 1				Db	271	LDNNKKLRLRPAGLAQHKYQVYVLHNNNNTSEVGQHDFCLPSYQTRKTSTAVSLISNPR	330
Best Local Similarity 45.1%; Score 89; DB 2; Length 354;				Qy	355	YWEMOPATFRCVLSRMSVOLGNF	377
Matches 178; Conservative 70; Mismatches 99; Indels 36; Gaps 5;				Db	331	YQJHPTFRCVFGRTSIOLGNY	353
1 MKEYVLLFLATCS-AKPFPSHSITALKNNMLKDMDTDDDDDDDEDNLSLFPTR 59				Db	211	PQGLPITSSELHDGNKIAKYDAASLKGMSNLSKUGLSFNSITVVENGSLANVPHLRELH	270
1 MKATLVLFLAQWSVGPFF--EQRLDFMLD -- ----- 31				Oy	295	LENNKIKTPGGLPEPKYLOLIFHNSTARVGVDNFCPTVKUKKTSALISLFNNPK	354
Qy 60 EPRSHFFPD --- LFPMPCCPGCQSYRSVHCSIDLGLTSVPTNIPFDTRMLDQNKKKE 115				Db	271	LDNNKKLRLRPAGLAQHKYQVYVLHNNNNTSEVGQHDFCLPSYQTRKTSTAVSLISNPR	330
Db 32 -EASGLIYPDPDNPLISMCPYRQCOHLRVYQCSIDLGLDKVWPWEFPDTTLDDQNKKTE 90				Qy	349	FNNPVKYWMQPATFRCVLSRMSVQL	374
Qy 116 IKENDFKGLTSVYGLLNNNNLTKIHPKAFLTKLKRLRPLNPKSLAEL 175				Db	331	DGNNULTRALPQEMYNCLRVAADISL	342
Db 91 IREGAKPNKLQKHTLILVNNKKSKISPEAKPLKLYSKHKEPEKPLKTQEL 150				Db	211	PQGLPITSSELHDGNKIAKYDAASLKGMSNLSKUGLSFNSITVVENGSLANVPHLRELH	270
Qy 176 RIHENKVKTQDFTKGMMALHLEMANSPLDNGTIEPAGFEGVY-VFHTRIAKLTSV 234				Oy	295	LENNKIKTPGGLPEPKYLOLIFHNSTARVGVDNFCPTVKUKKTSALISLFNNPK	354
Db 151 RHDNEETKLKSVNGLNRMIVLEGNNPLKNSTENGALQMGKGLGTRISDNITAI 210				Db	271	LDNNKKLRLRPAGLAQHKYQVYVLHNNNNTSEVGQHDFCLPSYQTRKTSTAVSLISNPR	330
Qy 235 PKGLPPTLLELHDYDYNKISTVELEDFKRYKELQRLGNNKTDIENGSLANIPRVRTH 294				Qy	355	YWEMOPATFRCVLSRMSVOLGNF	377
A;Title: Sequence and structural implications of a bovine corneal sulfate pro				Db	331	YQJHPTFRCVFGRTSIOLGNY	353
RESULT 12				Db	211	PQGLPITSSELHDGNKIAKYDAASLKGMSNLSKUGLSFNSITVVENGSLANVPHLRELH	270
A46743				Oy	295	LENNKIKTPGGLPEPKYLOLIFHNSTARVGVDNFCPTVKUKKTSALISLFNNPK	354
Iunican precursor - bovine				Db	271	LDNNKKLRLRPAGLAQHKYQVYVLHNNNNTSEVGQHDFCLPSYQTRKTSTAVSLISNPR	330
N;Alternate names: 37K keratan sulfate-linked protein; corneal keratan sulfate proteo				Qy	355	YWEMOPATFRCVLSRMSVOLGNF	377
C;Species: Bos primigenius taurus (cattle)				Db	331	YQJHPTFRCVFGRTSIOLGNY	353
C;Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 24-Sep-1999				Db	211	PQGLPITSSELHDGNKIAKYDAASLKGMSNLSKUGLSFNSITVVENGSLANVPHLRELH	270
C;Accession: A46743; A53379				Oy	295	LENNKIKTPGGLPEPKYLOLIFHNSTARVGVDNFCPTVKUKKTSALISLFNNPK	354
R;Funderburgh, J.L.; Funderburgh, M.L.; Brown, S.J.; Vergens, J.P.; Hassell, J.R.; Ma				Db	271	LDNNKKLRLRPAGLAQHKYQVYVLHNNNNTSEVGQHDFCLPSYQTRKTSTAVSLISNPR	330
J. Biol. Chem. 268, 11871-11880, 1993				Qy	355	YWEMOPATFRCVLSRMSVOLGNF	377

A; Reference number: A46743; MUID: 93280153; PMID: 8099356  
A; Molecule type: mRNA  
A; Residues: 1-342 <FUN>  
A; Cross-references: GB:L11063; NID:9163266.; PIDN:AAA30608.1; PID:9163267  
A; Experimental source: Cornea  
R; Funderburgh, J.L.; Conrad, G.W.  
J.; Biol. Chem. 265, 8297-8305, 1990  
A; Title: Isoforms of corneal keratan sulfate proteoglycan.  
A; Reference number: A35379; MUID: 90243714; PMID:2139877  
A; Accession: A35379  
A; Molecule type: protein  
A; Residues: 19-36 <FU2>  
A; Experimental source: cornea  
C; Superfamily: fibromodulin; leucine-rich alpha-2-glycoprotein repeat homology  
C; Keywords: glycoprotein; tandem repeat  
F; 1-18/Domain: signal sequence #status predicted <STG>  
F; 19-342/Domain: product: lumican #status predicted <MAT>  
F; 92,131,164,256/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 19-38%; Score 384; DB 2; Length 342;  
Best Local Similarity 32.28%; Pred. No. 6.3e-19;  
Matches 101; Conservative 68; Mismatches 123; Indels 22; Gaps 10;  
Db 139068

RESULT 14

66 FPFDLFF---PMCPFGCQC---YSRVVHCSIDLGLTSYPTNIPFDTRMLDLQNNKIKEIKE 118  
Db 29 FPOQAYGRSSPNCAPCEPEPSAMVCDELTLKSKVPM-VPGIKYLYRNQDIDDD 87

Qy - 119 NDFKGLTSLYGLLNNNNKL--TKTHPKAFLTTKRLRLYLSHNLSEIPLNPKSLAELR 176  
Db 88 KAFENVTDLQWLDHNLLENNSLKKGKVFSKQLQKKHLHINYNLNTESVGPLPKSLVDLQ 147

Qy - 177 THENVKRKQIQTDFKGMNALHVLEMSANPLDNG1IEPGAFEGV-TVFHIRIAEAKLTSPV 235  
Db 148 LTNNNKISKL-GSFDFGLVNLTIFLQHNLKEIDAVS-HALKGKLSLEYDLSFNQMKLP 204

236 KGLPPLTLELHLDYNIKISTVELEDFKRYKELORLGLGNKITDI-ENGSLANIPIVREITH 294  
Db 205 SGPLVSLTLYLDNNKISNIPDEFFKRSALQFLRLSHNEALASGVPONSFNSSLFLD 264

Qy 295 LENNNKKLKKPSGLPFLK--YLOI1FLHSNSIARVGVNDFCPTYPKMKSLYSAAISLNNP 352  
Db 265 LSYNKLRKSIPIVNNENJENYLEV----NELEKFVSKFCKILGPLSYSKIRHLQDGHN 319

Qy 353 VKYWMQMPATFRCV 366  
Db 320 ITQTSPLPPDMYECL 333

RESULT 15

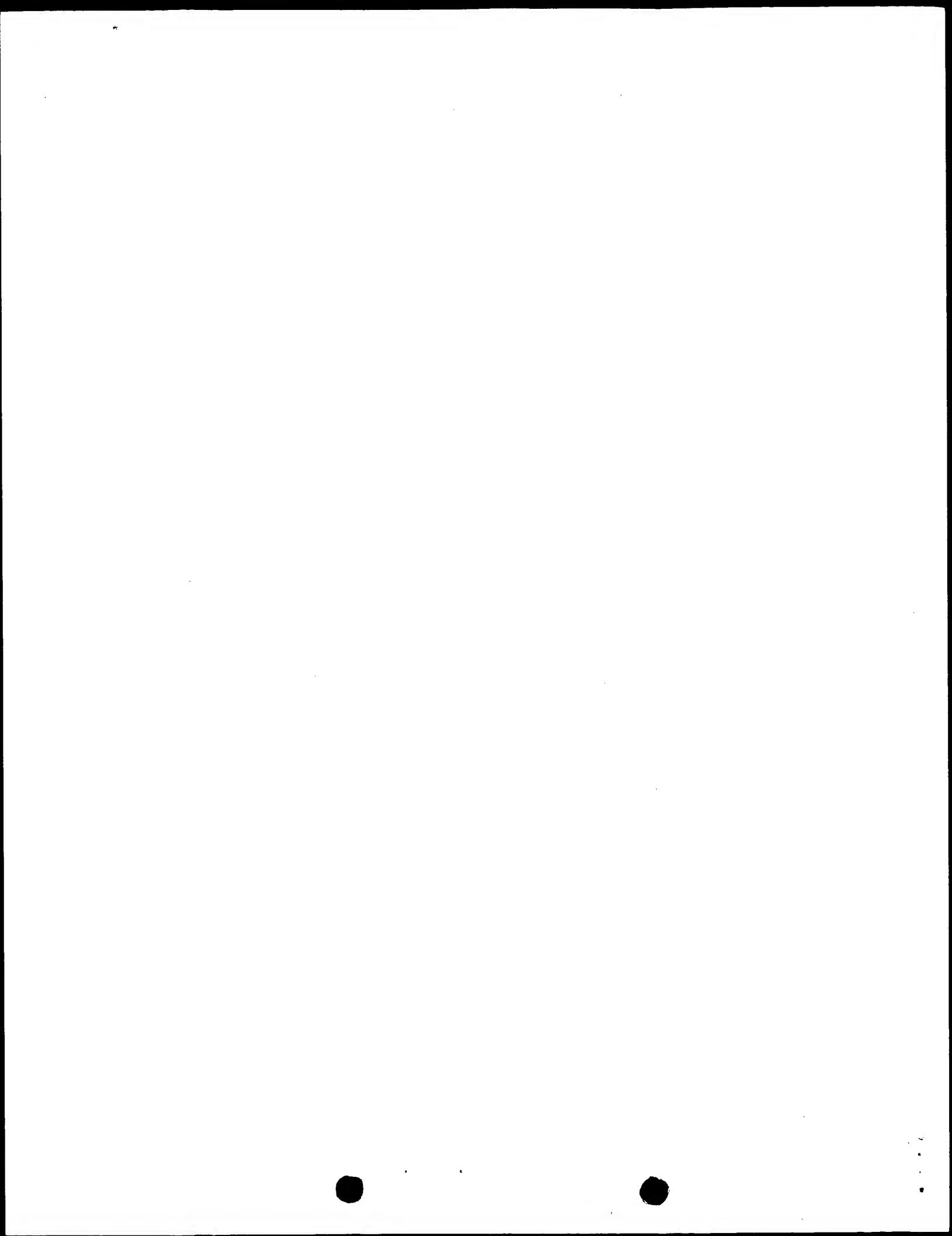
Q22284 Secretory interstitial proteoglycan precursor - rat  
Species: Rattus norvegicus (Norway rat)  
C; Date: 08-May-1995 #sequence\_revision 21-Jul-1995 #text\_change 24-Sep-1995  
R; Krull, N.B.  
submitted to the EMBL Data Library, January 1995  
A; Reference number: S52284  
A; Accession: S52284  
A; Status: preliminary  
A; Molecule type: mRNA  
A; Residues: 1-338 <KRU>  
A; Cross-references: EMBL:X84039; NID:9643023; PIDN:CAA58858.1; PID:9643024  
A; Superfamily: fibromodulin; leucine-rich alpha-2-glycoprotein repeat homology  
C; Best Local Similarity 32.28%; Pred. No. 1.6e-18;  
Matches 99; Conservative 67; Mismatches 123; Indels 18; Gaps 9;

Query Match 19.0%; Score 378; DB 2; Length 338;  
Db 199 SLEPMCPFGCQC---YSRVVHCSIDLGLTSYPTNIPFDTRMLDLQNNKITEKENDFKGAT 125  
Db 32 ELSPNCAPECNCPHSYPTAMYCDDLKLKSVPM-VPGIKYLYRNNOTDHIDRAFENVT 90

RESULT 15

S05390 fibromodulin precursor - bovine  
C; Alternate names: 59k collagen-binding matrix protein  
C; Species: Bos primigenius taurus (cattle)

Sequence Comparison Report						
Query Sequence		Target Sequence			Match Statistics	
C; Date: 22-Jan-1993 #sequence_revision 22-Jan-1993 #text_change 24-Sep-1999						
C; Accession: S03390; S26942; S06040						
R; Oldberg, A.; Antonsson, P.; Lindblom, K.; Heinegard, D.						
EMBO J., 8, 2601-2604, 1989						
A; Title: A collagen-binding 59-kd protein (fibromodulin) is structurally related to the						
A; Reference number: S05390; MUID:9060020; PMID:2531085						
A; Accession: S05390						
A; Molecular type: mRNA						
A; Residues: 1-375 <OL2>						
A; Cross-references: GB:X16485; EMBL:X16305; NID:9353; PIDN:CAA34503.1; PMID:g354						
A; Accession: S26942						
A; Molecular type: protein						
A; Residues: 90-105;150-199;274-281 <OL2>						
A; Plaas, A.H.K.; Nease, P.J.; Nivens, C.M.; Reiss, L.						
J. Biol. Chem. 265, 20634-20640, 1990						
A; Title: Identification of the keratan sulfate attachment sites on bovine fibromodulin.						
A; Reference number: A23663; MUID:9105619; PMID:2243109						
A; Content: annotation: keratan sulfate attachment environment: keratan sulfate can be covalently attached to N-acetylglicosamine at the asparagine residue: fibromodulin: leucine-rich alpha 2-glycoprotein repeat homology						
C; Keywords: cartilage; chondroitin sulfate proteoglycan; collagen binding; connective tissue; domain; signal sequence; status predicted <SIG>						
F; 1-18/Domain: signal sequence status predicted <SIG>						
F; 19-47/Domain: signal sequence status predicted <SIG>						
F; 38,42,45,47,50,53,62,64/Binding site: sulfate (Tyr) (covalent) #status predicted F;126,165,200,290/Binding site: carbohydrate (Asn) (covalent) #status experimental						
Query Match	17	2*	Score 342.5;	Length 375;		
Best Local Similarity	27.7%	Pred. No. 4.8e-16;				
Matches 103; Conservative 66; Mismatches 158; Indels 45; Gaps 13;						
Qy	6	LILFLALCSAK -	-PFPSPHITALKNNMMLKDMDTDDDDDD	-D D D D D D E N S I F -	56	
Db	6	I L L A G L S I S L S W A Q Y E E D D H W N F Q -	-F L R N Q Q S T Y D D Y D P P Y E P Y P F Y P G E E G P A Y A	63		
Qy	57	--P T R E P S H F F P D L F P M C P G F C Q --	-Y S R V H G S D L G L T S V P N I P F D T R M L D Q	109		
Db	*	. 64 Y G S P P Q P E R D -	-C P Q E C D P P N F P T A M Y C D N R N I K Y L P P - V P S R M K Y Y F Q	112		
Qy	110	N N K I K E I K E N D F K G L T S Y G L I N N K K L T -	-K I H P K A F L T K K U R R L Y L S H N Q L S E I P L N	167		
Db	.	. 113 N N Q T S I Q E G V F D N A T G L W I A L H G N Q I T S D K V G K V F S K R H L E R L Y L D H N L N T R I P S P	172			
Qy	168	L P K S I A E L R I H E N K V K K I Q K D T P K G M N A H V L E M S A N F D D N G I E P A F E G V - T V F H I R I	226			
Db	173	I P S R S L H D N O T S R P N N A L E G L E N T A L Y L H H N Q I Q E V G - - S S M K G L R S L I L D L	229			
Qy	227	A E A K L T S V P K G L P P T L L E H D Y N K I S T V E L E D K R Y K E L Q R L G N K N K I T D E N G S L A -	285			
Db	230	S Y N H L R K V D G L P S A L E Q Y L E H I N N V F S V P D S Y F R G S K L L Y V R L S H N S I L T - - N N G L A S	286			
Qy	286	- - N P V R Y R B H L E N N K K L K T P S G L P E I K Y L Q I T F L H S N S T A R V G V N D F C P T V P K M K K S L	342			
Db	287	N T F N S S S L E D L S Y N Q L Q K I P - - P V S T N L E N L Y L O G N R I N E F S I S S F C T V D V M N E S K	343			
Qy	343	Y S A T S L F N P P V K	354			
Db	344	L Q V O R D G N E I K	355			



Result No.	Score	Query	Match	Length	DB	ID	Description
1	1992	100.0	379	1	ASPN_HUMAN	Q9BXX1	Q9BXX1 homo sapien
2	1799	90.3	373	1	ASPN_MOUSE	Q99mg4	mus musculi
3	1044	52.4	568	1	PGSL_XENLA	Q9ib75	xenopus lae
4	1040.5	52.2	369	1	PGSL_RAT	P47853	rattus norv
5	1037.5	52.1	369	1	PGSL_MOUSE	P28653	mus musculi
6	1034	51.9	368	1	PGSL_HUMAN	P21810	homo sapien
7	1028.5	51.6	372	1	PGSL_HORSE	Q46403	equus cabal
8	1021.5	51.6	369	1	PGSL_SHEEP	Q46390	ovis aries
9	1021.5	51.3	369	1	PGSL_CANFA	Q02678	canis famil
10	1019.5	51.2	369	1	PGSL_BOVIN	P21809	bos taurinus
11	974	48.9	356	1	PGS2_CORTJA	Q9de68	coturnix co
12	969	48.6	357	1	PGS2_CHICK	P28675	gallus galli
13	967	48.5	360	1	PGS2_PIG	Q9xsd9	sus scrofa
14	963.5	48.4	359	1	PGS2_HUMAN	P07855	homo sapien
15	947	47.5	360	1	PGS2_SHEEP	Q9te2	ovis aries
16	943	47.3	360	1	PGS2_BOVIN	P21793	bos taurinus
17	941	47.2	360	1	PGS2_RABBT	P28888	oryctolagus cuniculus
18	939	47.1	360	1	PGS2_HORSE	Q46542	equus cabal
19	934	46.9	360	1	PGS2_CANFA	Q29393	canis familiaris
20	907.5	45.6	354	1	PGS2_MOUSE	P28654	mus musculi
21	899	45.1	354	1	PGS2_RAT	Q01129	rattus norvegicus
22	641	32.2	272	1	PGS1_PIG	Q99kq6	sus scrofa
23	442	22.2	135	1	PGS1_RABBT	Q46377	oryctolagus cuniculus
24	399.5	20.1	343	1	LUM_CHICK	P511990	gallus gallus
25	398.5	20.0	343	1	LUM_COTJA	Q9de67	coturnix coturnix coquettiae
26	390	19.6	338	1	LUM_HUMAN	P511884	homo sapien
27	384	19.3	342	1	LUM_BOVIN	P05443	bos taurinus
28	383	19.2	338	1	LUM_MOUSE	P511885	mus musculus
29	378	19.0	338	1	LUM_RAT	P511886	rattus norvegicus
30	366	18.4	421	1	OMD_HUMAN	Q99n93	homo sapien
31	363.5	18.2	646	1	FLR_HUMAN	Q9zul2	homo sapien
32	361.5	18.1	381	1	PRLP_BOVIN	Q99kn8	bos taurinus
33	358.5	18.0	382	1	PRLP_HUMAN	P511888	homo sapien

ALIGMENTS							
RESULT 1							
ASPN_HUMAN							
ID	ASPN_HUMAN	STANDARD;					PRT;
AC	Q9BXX1;	Q96LDD0;	Q96K79;				
DT	15-JUN-2002	(Rel. 41,	Created)				
DT	15-JUN-2002	(Rel. 41,	Last sequence update)				
DT	15-JUN-2002	(Rel. 41,	Last annotation update)				
DE	Asporin precursor	(Periodontal ligament associated protein-1)					
DE	1).						
GN	ASPN OR PLAP1.						
OS	Homo sapiens (Human).						
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;						
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.						
OX	NCBI_TAXID=9606;						
RN	SEQUENCE FROM N.A., PARTIAL SEQUENCE, CARBOHYDRATE-LINKAGE SITE						
RC	ASN-281, MASS SPECTROMETRY, AND POLYMORPHISM OF POLY-ASP REGION						
RC	TISSUE-Cartilage;						
RX	MEDLINE=2119276; PubMed=11152692;						
RA	Lorenzo P., Asperberg A., Oennerfjord P., Bayliss M.T., Neame P.J., Heinegaard D.						
RT	"Identification and characterization of asporin. A novel member of the leucine-rich repeat protein family closely related to decorin and biglycan."						
RT	J. Biol. Chem. 276:12201-12211 (2001).						
RP	SEQUENCE FROM N.A.						
RX	MEDLINE=2147263; PubMed=11587855;						
RA	Yamada S., Murakami S., Matoba R., Ozawa Y., Yokokoji T., Nakahira Y., Ikezawa K., Takayama S.-I., Matsubara K., Okada H.;						
RT	"Expression profile of active genes in human periodontal ligament and isolation of PLAP-1, a novel SLRP family gene.";						
RL	Gene 275:279-286 (2001).						
RN	SEQUENCE OF 1-242 FROM N.A.						
RC	TISSUE=Embryo;						
RA	Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y., Nishikawa T., Nagai K., Shiratori A., Sudo H., Hidemoto R., Zhou W., de Crombrugghe B., Hoeoeik M., Mayne R.						
RA	Wagatsuma M., Hosoi T., Kaku Y., Kodaira H., Kondo H., Sugawara M., Takahashi M., Ishida S., Murakawa K., Ono Y., Takiguchi S., Watamabe J., Kimura K., Murakami K., Ishii S., Kawai Y., Yamamoto J., Wakamatsu A., Nakamura K., Nagahari K., Masuno Y.;						
RA	Ninomiya K., Iwayanagi T., NEDO human cDNA sequencing project.;						
RL	Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.						
[4]	PARTIAL SEQUENCE FROM N.A.						
RX	PubMed=11152695;						
RA	Henry S.P., Takano M., Boyd T.C., Mayne P.M., Eberspaecher H., Zhou W., de Crombrugghe B., Hoeoeik M., Mayne R.						
RT	"Expression pattern and gene characterization of asporin. A newly discovered member of the leucine-rich repeat protein family.";						
RL	J. Biol. Chem. 276:12212-12221 (2001).						
CC	- - - SUBCELLULAR LOCATION: Extracellular matrix.						
CC	- - - TISSUE SPECIFICITY: Higher levels in osteoarthritic articular cartilage, aorta, uterus. Moderate expression in small intestine,						

CC heart, liver, bladder, ovary, stomach, and in the adrenal,  
 CC thyroid, and mammary glands. Low expression in trachea, bone  
 CC marrow, and lung.  
 CC -!- PTM: Does not contain a serine/glycine dipeptide sequence required  
 CC for the assembly of O-linked glycosaminoglycan and is probably  
 CC not a proteoglycan. The N-linked glycan at Asp-281 is composed of  
 CC variable structures of GlcNAc, mannose, fucose, HexNAc and hexose.  
 CC -!- CC POLYMORPHISM: The poly-Asp region of ASPN is polymorphic and  
 CC ranges at least from 11 to 17 Asp  
 CC -!- SIMILARITY: BELONGS TO THE SMALL LEUCINE-RICH PROTEOGLYCANS  
 CC (SLRPs) FAMILY. CLASS: 12 LEUCINE-RICH REPEATS (LRR).  
 CC -!- SIMILARITY: CONTAINS 12 LEUCINE-RICH REPEATS (LRR).  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires license agreement. (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to license@sb-sib.ch).  
 CC

DR EMBL; AF316824; AAK5161\_1; -;  
 EMBL; AY029191; AAK31800\_1; -;  
 EMBL; AK027359; BAB5060\_1; -;  
 DR Genbank; HGNC:14872; ASPN  
 DR InterPro; IPR001611; LRR.  
 DR InterPro; IPR000372; LRR\_Nterm.  
 DR InterPro; IPR003592; LRR\_out.  
 DR Pfam; PF00560; LRR; 8.  
 DR Pfam; PF01462; LRRNT; 1.  
 DR SMART; SM00370; LRR; 3.  
 DR SMART; SM00013; LRRNF; 1.  
 DR SMART; SM00369; LRR\_TYP; 7.  
 KW Glycoprotein; Extracellular matrix; Signal; Repeat;  
 KW Leucine-rich repeat; Polymorphism; Triplet repeat expansion.  
 FT SIGNAL 1 14 POTENTIAL.  
 FT PROPEP 15 32 ASPORN.  
 FT CHAIN 33 379  
 FT DOMAIN 76 89 CYS-RICH.  
 FT REPEAT 95 115 LRR-S 1.  
 FT REPEAT 116 139 LRR-T 1.  
 FT REPEAT 140 163 LRR-T 2.  
 FT REPEAT 164 184 LRR-S 2.  
 FT REPEAT 185 208 LRR-T 3.  
 FT REPEAT 209 233 LRR-T 4.  
 FT REPEAT 234 254 LRR-S 3.  
 FT REPEAT 255 278 LRR-T 5.  
 FT REPEAT 279 302 LRR-T 6.  
 FT REPEAT 303 325 LRR-S 4.  
 FT REPEAT 326 355 LRR-T 7.  
 FT REPEAT 356 379 LRR-T 8.  
 DISULFID 332 365 BY SIMILARITY.  
 DOMAIN 36 52 POLY-ASP.  
 FT CARBOHYD 54 54 O-LINKED (BY SIMILARITY).  
 FT CARBOHYD 281 281 N-LINKED (GLCNAC. . . ).  
 FT CONFLICT 237 242 GLPPTP -> DNLPSPF ( IN REF. 3 ).  
 SQ SEQUENCE 379 AA; 43302 MW; 78663852E26A6D CRC64;  
 SEQUENCE [REDACTED] 100.0%; Score 1992; DB 1; Length 379;  
 Best Local Similarity 100.0%; Pred. No. 1. 2e-13.0;  
 Matches 379; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKEYVLLFLAUSAKPEFPSSPHIAALKNMKLKDMDTDDDDDDDDDDNSLEPTRE 60  
 1 MKEYVLLFLAUSAKPEFPSSPHIAALKNMKLKDMDTDDDDDDDDDDNSLEPTRE 60  
 DB 181 KVKKIQRDTFGMMNLHVLEMSANPLDNNGTEPGAFEGTVFHIRIAEAALKTSVPKGLLP 240  
 DB 181 KVKKIQRDTFGMMNLHVLEMSANPLDNNGTEPGAFEGTVFHIRIAEAALKTSVPKGLLP 240  
 QY 241 TLLELHDYKNTSTVELEDFKRYKEQLQRGLGNNKTIDTNGSLANIPTREIHLENKL 300  
 QY 241 TLLELHDYKNTSTVELEDFKRYKEQLQRGLGNNKTIDTNGSLANIPTREIHLENKL 300  
 DB 241 KKIPLSGPLPEKYLQIIFLHSNSIARYGVNDFCPVPKMKRSLYSAISLNPNPKWEMQP 360  
 DB 301 KKIPLSGPLPEKYLQIIFLHSNSIARYGVNDFCPVPKMKRSLYSAISLNPNPKWEMQP 360  
 QY 301 KKIPLSGPLPEKYLQIIFLHSNSIARYGVNDFCPVPKMKRSLYSAISLNPNPKWEMQP 360  
 DB 301 KKIPLSGPLPEKYLQIIFLHSNSIARYGVNDFCPVPKMKRSLYSAISLNPNPKWEMQP 360  
 QY 361 ATFRCVLSRMSVQLGNFGM 379  
 DB 361 ATFRCVLSRMSVQLGNFGM 379  
 RESULT 2  
 ASN\_MOUSE STANDARD PRT: 373 AA.  
 ID ASN\_MOUSE STANDARD PRT: 373 AA.  
 AC Q99MQ4; Q9D6A2;  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Asporin precursor.  
 DR ASNIN.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.; TISSUE=Aorta;  
 RC STRAIN=C57BL/6J; TISSUE=Skin;  
 RX MEDLINE=21192276; PubMed=11152692;  
 J. Biol. Chem. 276:12201-12211(2001).  
 RA Lorenzini P., Asperberg A., Oennertjord P., Bayliss M.T., Neame P.J., Heinegaard D.; "Identification and Characterization of Asporin. A novel member of the leucine-rich repeat protein family closely related to decorin and biglycan"; J. Biol. Chem. 276:12201-12211(2001).  
 RN [2]  
 RP SEQUENCE FROM N.A.; AND DEVELOPMENTAL STAGE.  
 RA Henry S.P., Takano M., Boyd T.C., Mayne P.M., Eberspacher H., Zhou W., de Crombrugghe B., Hoeck M., Mayne R.; "Expression pattern and gene characterization of asporin. A newly discovered member of the leucine-rich repeat protein family.";  
 RA Saito R., Okazaki Y., Gojobori T., Bono H., Kusakawa T., Saito R., RL J. Biol. Chem. 276:12201-12211(2001).  
 RN [3]  
 RP SEQUENCE FROM N.A.; STRAIN=C57BL/6J; TISSUE=Skin;  
 RC MEDLINE=21085660; PubMed=11217851;  
 RA Kawai J., Shinagawa A., Shibusawa K., Yoshino M., Itoh M., Ishii Y., Arakawa T., Har A., Fukunishi Y., Konno H., Adachi J., Fukuda S., Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamamoto T., Saito T., Okada T., Furuno M., Tomita M., Wagner L., Washio T., Kadda K., Matsuda H.A., Asiburner M., Baralov S., Casavant T., Fleschmann W., Gaasterland T., Gissi C., King B., Kochiwa H., Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J., Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T., Sakai K., Okido T., Baldarelli R., Barsh G., Blate J., Botelli D., Bojunga N., carinici P., de Bonaldo M.F., Bronstein M.J., Bult C., Fletcher C., Gariboldi M., Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H., Lyons P., Marchionni L., Mashima J., Mazzarelli J., Monbaerts P., Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N., Sato K., Schoenbach C., Seya T., Shibusawa Y., Storch K., Suzuki H., Toyooka K., Wang K.H., Whittaker C., Wilming L., Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohutsu S., Hayashizaki Y.; "Functional annotation of a full-length mouse cDNA collection.";

Nature 409:685-690(2001).

-!- SUBCELLULAR LOCATION: Extracellular matrix (By similarity).

-!- TISSUE SPECIFICITY: Higher expression in heart, also detected in kidney, stomach, testes, and skin but only weakly in lung, skeletal muscle, small intestine, and thymus.

-!- DEVELOPMENTAL STAGE: At 12.5 dpc, it is present in the mandibular as well as maxillary components of the first branchial arch. Also detected in the thoracic body wall adjacent to the heart. At 13.5 dpc, it is detected in the mesenchyme lateral to Meckel's cartilage. Pronounced expression is observed in the perichondrium of the humerus, ribs, and scapula. At 14.5 dpc, it is detected in the mesenchymal condensations lateral to Meckel's cartilage, in the perichondrium surrounding the central cartilaginous elements of the vertebra and also in the dermal mesenchyme. At 15.5 dpc, it is expressed in the perichondrium/periosteum of the long bones (i.e. femur, tibia, and fibula). Some of the flat bones at the base of the skull (i.e. sphenoid bone), ribs, clavicle, and vertebrae. Also detected in the intramembranous bones of the maxilla and mandible (alveolar bone) and a strong expression is observed in sagittal sections of the subcutaneous muscles or panniculus carnosus of the thorax, trunk, and head/neck (platysma muscle) region. Very little expression is detected in the major parenchymal organs (with the exception of the large bronchi of the lung). Its expression is prominent in the developing mouse skeleton, particularly in the perichondrium/periosteum of cartilage/bone, and is also found in other specialized connective tissues such as tendon, sclera, the connective tissue sheath surrounding muscle and dermis. In the sciara of the eye it is first detected at 15.5 dpc and stronger expression was detected at 17.5 dpc.

-!- SIMILARITY: BELONGS TO THE SMALL LEUCINE-RICH PROTEOGLYCANS (SLRPs) FAMILY. CLASS I SUBFAMILY.

-!- SIMILARITY: CONTAINS 12 LEUCINE-RICH REPEATS (LRR).

-!- CAUTION: Ref. 3 sequence differs from that shown due to a stop

FT	REPEAT	350	373	LRR-T 8.			
FT	DOMAIN	36	46	POLY-ASP.			
FT	DISULFID	326	359	BY SIMILARITY.			
FT	CARBOHYD	48	48	O-LINKED (BY SIMILARITY).			
FT	CARBOHYD	275	275	N-LINKED (GHCAC- : -). (POTENTIAL).			
SQ	SEQUENCE	373 AA:	42572 MW:	A50CA4C82AABCF25 CRC64;			
Query Match	Best Local Similarity	90.3%	Score 90.3%	Length 373;			
Matches	Local Similarity	90.2%	Pred. No. 2	Length 373;			
Matches	Conservative	17	Mismatches	14;			
			Indels	6;			
			Gaps	1;			
QY	1	MKEYVLLIAKCSAKPPESSPHIALALKNMKLKDMEETDDDDDDDDDDNSLEPLTE 60					
Db	1	MKEYVMILLAVCSAKPPESSPHIALALKNMKLKDMEETDDDDDDDD----DNSLFPTRKE 54					
QY	61	PRSHFFPPDLPMPMCPEFCQCYSRVICSDLSLTSVPTNPFDTRMLDNQNKKIEKEND 120					
Db	55	PYNPFEEFLDFPTCPFPGCQCYSRVICSDLSLTSVPTNPFDTRMLDNQNKKIEKEND 114					
Qy	121	PKGLTSLYGLILNNNNKLTKIHKPAFLTKIHKRLYLPSLAELRTHEN 180					
Db	115	FKGLTSLYALILNNNNKLTKIHKRTFLTKIHKRLYLPSLAELRHDN 174					
Qy	181	KYKKIQDKTFKGMMNLHVLEMSANPLDNNNGEPGAEGIVVPHIRIAEAKLTSPVKGLP 240					
Db	175	KYKKIQDKTFKGMMNLHVLEMSANPLENNNGEPGAEGIVVPHIRIAEAKLTSFKGLP 234					
Qy	241	TLLELHDYDKNKTSTVEEDFKRYKEQLQRGLGNNKTTDINGSLANIPTYREIHLENKL 300					
Db	235	TLLELHDYDKNKTSTVEEDFKRYKEQLQRGLGNNKTTDINGTFANIPTYREIHLENKL 294					
Qy	301	KKTPSGLPLPEKLYLQITFLHNSISIARYGVNDFCPPTVPKMKSLSAISLFNNPVKYWEMQP 360					
Db	295	KKTPSGLPLPEKLYLQITFLHNSISIARYGVNDFCPPTVPKMKSLSAISLFNNPVKYWEIQP 354					
Qy	361	ATFRCVLSRMSVOLGNFG 378					
Db	355	ATFRCVLRMSVOLGNVG 372					
RESULT 3							
PGS1_XENLA							
ID	PGS1_XENLA	STANDARD:	PRT:	368 AA.			
AC	Q91IB75;						
DT	15-JUN-2002	(Rel. 41, Created)					
DT	15-JUN-2002	(Rel. 41, Last sequence update)					
DT	15-JUN-2002	(Rel. 41, Last annotation update)					
DE	Biglycan precursor.						
GN	BGN.						
OS	Xenopus laevis (African clawed frog).						
OC	Amphibia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;						
OC	Xenopoda; Batracchia; Anura; Mesobatrachia; Pipoidea; Pipidae;						
OC	Xenopodinae; Xenopus.						
OX	NCBI_TAXID-8355;						
RN	[1]	SEQUENCE FROM N.A.					
RP	Goto T., Kubota H.Y.;						
RA	*CDNA of biglycan of <i>Xenopus laevis</i> ;						
RT	Submitted (JAN-2000) to the EMBL/GenBank/DDJB databases.						
RL	- - FUNCTION: May be involved in collagen fiber assembly (By similarity).						
CC	- - SUBCELLULAR LOCATION: Extracellular matrix (By similarity).						
CC	- - SIMILARITY: BELONGS TO THE SMALL LEUCINE-RICH PROTEOGLYCANS (SURPs) FAMILY. CLASS 1 SUBFAMILY.						
CC	- - SIMILARITY: CONTAINS 12 LEUCINE-RICH REPEATS (LRR).						
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> ) or send an email to license@isb-sib.ch						





Db	110 DFKGQHQLAVLNKNSKTHEKAFPSLRLKLYTSKNHLVEIPPNLSSVELRIHD 169	RL Biochem. J. 262:823-827 (1989).
Qy	180 NKVKIQKDTPKGMMALHYEMSLANPLDNGIIPGAEGFVTVWHIRIAEAKLTSPVKGLP 239	RN [6] SEQUENCE OF 38-66.
Db	170 NRIRRVPKGSGLRRNMCLIEGGNPLENSGFFPGAFDGLKLYRSEAKLTGIPKDLP 229	RX MEDLINE=87250639; PubMed=3597437;
Qy	240 PTLLELHDYWKNTISTVELEDKRYKELQRIGLGNKTKDIENGLANIPRVTEIHLNNK 299	RA Fisher L.W., Hawkins G.R., Tuross N., Termino J.D.; RT Purification and partial characterization of small proteoglycans I and II, bone sialoproteins I and II, and osteonectin from the mineral compartment of developing human bone."; RT J. Biol. Chem. 262:9702-9708(1987).
Db	230 ETLNEELHDINKNQAEDELRLRLLGLQHQNIRMENGLSFLPTLRLHDLNK 289	RL RN [7]
Qy	300 LKIPSPGLPELKYLQTFILHSNSLSTARGVYNDFCPVPMKKSLSATISLNFNEVKYWMQ 359	RP TISSUE-SKIN; SEQUENCE FROM N.A.
Db	290 LSRPVAGLPLDLKLQVVLHSNNLTKVGINDFCPMGFVWKRATYNGTISLFNNEVPEVQ 349	RC MEDLINE=951187185; PubMed=7881444;
Qy	360 PATFRCVLSRMSVOLGNF 377	RA Just W., Rau W., Muller C., Geerkens C., Vogel W.; RT "Dinucleotide repeat polymorphism at the human biglycan (BGN) locus";
Db	350 PATFRCVTDRLAIOFGNY 367	RL Hun. Mol. Genet. 3:2268-2268(1994).
<b>RESULT 6</b>		CC -!- FUNCTION: May be involved in collagen fiber assembly (By similarity).
PS51 HUMAN		CC -!- SUBUNIT: Forms a ternary complex with MFAP2 and ELN (By similarity).
PGS1_HUMAN		CC -!- TISSUE SPECIFICITY: Found in several connective tissues, especially in articular cartilages.
DT 01-JAN-1990 (Rel. 13, Created)		CC -!- PTM: The two attached glycosaminoglycan chains can be either chondroitin sulfate or dermatan sulfate (By similarity).
DT 01-APR-1993 (Rel. 25, Last sequence update)		CC -!- SIMILARITY: BELONGS TO THE SMALL LEUCINE-RICH PROTEOGLYCANS (SLRPs) FAMILY. CLASS 1 SUBFAMILY.
DT 15-JUN-2002 (Rel. 41, Last annotation update)		CC -!- SIMILARITY: CONTAINS 12 LEUCINE-RICH REPEATS (LRR).
DE Biglycan precursor (Bone/cartilage proteoglycan 1) (PG-S1).		CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See <a href="http://www.ebi.ac.uk/announce/">http://www.ebi.ac.uk/announce/</a> or send an email to license@ebi.ac.uk).
GN BGN		CC
OS Homo sapiens (Human)		CC
OC Mammalia		CC
OC Metazoa		CC
OC Buthidae		CC
OC Primates		CC
OC Catarrhini		CC
OC Hominidae		CC
OX NCBI_TaxID=9606;		CC
RN 11		CC
RP SEQUENCE FROM N.A.		CC
RC TISSUE-BONE		CC
RX MEDLINE=89174714; PubMed=2647739;		DR EMBL: J04599; AAA6009_1;
RT Fisher L.W., Termino J.D., Young M.F.;		DR EMBL: M65153; AAAS287_1; ALT_SEQ.
RT "Reduced protein sequence of bone small proteoglycan I (biglycan)		DR EMBL: M65152; AAA5287_1; JOINED.
RT shows homology with proteoglycan II (decorin) and several		DR EMBL: U82695; AAF35530_1;
RT nonconnective tissue proteins in a variety of species.";		DR EMBL: BC00416; AAH02416_1;
RT J. Biol. Chem. 264:4571-4576(1989).		DR EMBL: BC00424; AAH04244_1;
RN [2]		DR EMBL: U11686; AAC50117_1;
RN [3]		DR PIR: A28457; AA28457;
RN [4]		DR PIR: A32458; A32458;
RN [5]		DR PIR: A40757; A40757;
RN [6]		DR PIR: S05639; S05639;
RN [7]		DR Genew; HGNC:1044; BGN.
RN [8]		DR MIM: 301870; -;
RN [9]		DR InterPro: IPR001611; LRR.
RN [10]		DR InterPro: IPR00372; LRR_Nterm.
RN [11]		DR InterPro: IPR003592; LRR_out.
RN [12]		DR InterPro: IPR003591; LRR_TYP.
RN [13]		DR Pfam: PF00560; LRR; 8 .
RN [14]		DR Pfam: PF01462; LRRNT; 1 .
RN [15]		DR SMART: SM00370; LRR; 3 .
RN [16]		DR SMART: SM00133; LRRNT; 1 .
RN [17]		DR Glycoprotein; Extracellular matrix; Proteoglycan; Repeat;
RN [18]		KW Leucine-rich repeat; Signal. POTENTIAL.
RN [19]		FT SIGNAL 1 19
RN [20]		FT PROPEP 20 37
RN [21]		FT CHAIN 38 368
RN [22]		FT DOMAIN 63 76
RN [23]		FT REPEAT 82 102
RN [24]		FT REPEAT 103 126
RN [25]		FT REPEAT 127 150
RN [26]		FT REPEAT 151 171
RN [27]		FT REPEAT 172 195
RN [28]		FT REPEAT 196 220
RN [29]		FT REPEAT 221 241
RN [30]		FT REPEAT 242 251
RN [31]		FT REPEAT 252 260
RN [32]		FT REPEAT 261 270
RN [33]		FT REPEAT 271 280
RN [34]		FT REPEAT 281 290
RN [35]		FT REPEAT 291 300
RN [36]		FT REPEAT 301 310
RN [37]		FT REPEAT 311 320
RN [38]		FT REPEAT 321 330
RN [39]		FT REPEAT 331 340
RN [40]		FT REPEAT 341 350
RN [41]		FT REPEAT 351 360
RN [42]		FT REPEAT 361 370
RN [43]		FT REPEAT 371 380
RN [44]		FT REPEAT 381 390
RN [45]		FT REPEAT 391 400
RN [46]		FT REPEAT 401 410
RN [47]		FT REPEAT 411 420
RN [48]		FT REPEAT 421 430
RN [49]		FT REPEAT 431 440
RN [50]		FT REPEAT 441 450
RN [51]		FT REPEAT 451 460
RN [52]		FT REPEAT 461 470
RN [53]		FT REPEAT 471 480
RN [54]		FT REPEAT 481 490
RN [55]		FT REPEAT 491 500
RN [56]		FT REPEAT 501 510
RN [57]		FT REPEAT 511 520
RN [58]		FT REPEAT 521 530
RN [59]		FT REPEAT 531 540
RN [60]		FT REPEAT 541 550
RN [61]		FT REPEAT 551 560
RN [62]		FT REPEAT 561 570
RN [63]		FT REPEAT 571 580
RN [64]		FT REPEAT 581 590
RN [65]		FT REPEAT 591 600
RN [66]		FT REPEAT 601 610
RN [67]		FT REPEAT 611 620
RN [68]		FT REPEAT 621 630
RN [69]		FT REPEAT 631 640
RN [70]		FT REPEAT 641 650
RN [71]		FT REPEAT 651 660
RN [72]		FT REPEAT 661 670
RN [73]		FT REPEAT 671 680
RN [74]		FT REPEAT 681 690
RN [75]		FT REPEAT 691 700
RN [76]		FT REPEAT 701 710
RN [77]		FT REPEAT 711 720
RN [78]		FT REPEAT 721 730
RN [79]		FT REPEAT 731 740
RN [80]		FT REPEAT 741 750
RN [81]		FT REPEAT 751 760
RN [82]		FT REPEAT 761 770
RN [83]		FT REPEAT 771 780
RN [84]		FT REPEAT 781 790
RN [85]		FT REPEAT 791 800
RN [86]		FT REPEAT 801 810
RN [87]		FT REPEAT 811 820
RN [88]		FT REPEAT 821 830
RN [89]		FT REPEAT 831 840
RN [90]		FT REPEAT 841 850
RN [91]		FT REPEAT 851 860
RN [92]		FT REPEAT 861 870
RN [93]		FT REPEAT 871 880
RN [94]		FT REPEAT 881 890
RN [95]		FT REPEAT 891 900
RN [96]		FT REPEAT 901 910
RN [97]		FT REPEAT 911 920
RN [98]		FT REPEAT 921 930
RN [99]		FT REPEAT 931 940
RN [100]		FT REPEAT 941 950
RN [101]		FT REPEAT 951 960
RN [102]		FT REPEAT 961 970
RN [103]		FT REPEAT 971 980
RN [104]		FT REPEAT 981 990
RN [105]		FT REPEAT 991 1000
RN [106]		FT REPEAT 1001 1010
RN [107]		FT REPEAT 1011 1020
RN [108]		FT REPEAT 1021 1030
RN [109]		FT REPEAT 1031 1040
RN [110]		FT REPEAT 1041 1050
RN [111]		FT REPEAT 1051 1060
RN [112]		FT REPEAT 1061 1070
RN [113]		FT REPEAT 1071 1080
RN [114]		FT REPEAT 1081 1090
RN [115]		FT REPEAT 1091 1100
RN [116]		FT REPEAT 1101 1110
RN [117]		FT REPEAT 1111 1120
RN [118]		FT REPEAT 1121 1130
RN [119]		FT REPEAT 1131 1140
RN [120]		FT REPEAT 1141 1150
RN [121]		FT REPEAT 1151 1160
RN [122]		FT REPEAT 1161 1170
RN [123]		FT REPEAT 1171 1180
RN [124]		FT REPEAT 1181 1190
RN [125]		FT REPEAT 1191 1200
RN [126]		FT REPEAT 1201 1210
RN [127]		FT REPEAT 1211 1220
RN [128]		FT REPEAT 1221 1230
RN [129]		FT REPEAT 1231 1240
RN [130]		FT REPEAT 1241 1250
RN [131]		FT REPEAT 1251 1260
RN [132]		FT REPEAT 1261 1270
RN [133]		FT REPEAT 1271 1280
RN [134]		FT REPEAT 1281 1290
RN [135]		FT REPEAT 1291 1300
RN [136]		FT REPEAT 1301 1310
RN [137]		FT REPEAT 1311 1320
RN [138]		FT REPEAT 1321 1330
RN [139]		FT REPEAT 1331 1340
RN [140]		FT REPEAT 1341 1350
RN [141]		FT REPEAT 1351 1360
RN [142]		FT REPEAT 1361 1370
RN [143]		FT REPEAT 1371 1380
RN [144]		FT REPEAT 1381 1390
RN [145]		FT REPEAT 1391 1400
RN [146]		FT REPEAT 1401 1410
RN [147]		FT REPEAT 1411 1420
RN [148]		FT REPEAT 1421 1430
RN [149]		FT REPEAT 1431 1440
RN [150]		FT REPEAT 1441 1450
RN [151]		FT REPEAT 1451 1460
RN [152]		FT REPEAT 1461 1470
RN [153]		FT REPEAT 1471 1480
RN [154]		FT REPEAT 1481 1490
RN [155]		FT REPEAT 1491 1500
RN [156]		FT REPEAT 1501 1510
RN [157]		FT REPEAT 1511 1520
RN [158]		FT REPEAT 1521 1530
RN [159]		FT REPEAT 1531 1540
RN [160]		FT REPEAT 1541 1550
RN [161]		FT REPEAT 1551 1560
RN [162]		FT REPEAT 1561 1570
RN [163]		FT REPEAT 1571 1580
RN [164]		FT REPEAT 1581 1590
RN [165]		FT REPEAT 1591 1600
RN [166]		FT REPEAT 1601 1610
RN [167]		FT REPEAT 1611 1620
RN [168]		FT REPEAT 1621 1630
RN [169]		FT REPEAT 1631 1640
RN [170]		FT REPEAT 1641 1650
RN [171]		FT REPEAT 1651 1660
RN [172]		FT REPEAT 1661 1670
RN [173]		FT REPEAT 1671 1680
RN [174]		FT REPEAT 1681 1690
RN [175]		FT REPEAT 1691 1700
RN [176]		FT REPEAT 1701 1710
RN [177]		FT REPEAT 1711 1720
RN [178]		FT REPEAT 1721 1730
RN [179]		FT REPEAT 1731 1740
RN [180]		FT REPEAT 1741 1750



**RESULT 8**

ID	PGS1_SHEEP	STANDARD;	PRT;	369 AA.
046390	046390			Q46390- (Rel. 39, Created)
D1	30-MAY-2000	(Rel. 39, Last sequence update)		
DT	30-MAY-2000	(Rel. 39, Last annotation update)		
DE	15-JUN-2002	(Rel. 41, Last annotation update)		
GN	Ovis	Biglycan precursor (Bone/cartilage proteoglycan 1) (PG-S1),		
OS	Ovis aries (Sheep).	Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Eukaryota; Metazoa;	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;		
OC	Bovidae;	Caprinae; ovis.		
OX	NCBI_TAXID:9940;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE-Choroid Plexus;			
RA	Bruett L., Clements J.E.;			
CC	Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases. (By similarity).			
-I-	FUNCTION: May be involved in collagen fiber assembly (By similarity).			
CC	-I- SUBUNIT: Forms a ternary complex with MFAP2 and FN (By similarity).			
CC	-I- SUBCELLULAR LOCATION: Extracellular matrix (By similarity).			
CC	-I- TISSUE SPECIFICITY: Found in several connective tissues, specially in articular cartilages.			
CC	-I- PTM: The two attached glycosaminoglycan chains can be either chondroitin sulfate or dermatan sulfate (By similarity).			
CC	-I- SIMILARITY: BELONGS TO THE SMALL LEUCINE-RICH PROTEOGLYCANS (SRP5) FAMILY. CLASS I SUBFAMILY.			
CC	-I- SIMILARITY: CONTAINS 12 LEUCINE-RICH REPEATS (LRR).			

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL: AF034842; AAB87988.1; -

CC DR InterPro: IPR001611; LRR\_Nterm.

CC DR InterPro: IPR000372; LRR\_out.

CC DR InterPro: IPR003531; LRR\_Eip.

CC DR Pfam: PF00560; LRR\_7.

CC DR Pfam: PF01462; LRRNT\_1.

CC DR SMART; SM000370; LRR\_3.

CC DR SMART; SM00013; LRRNT\_1.

CC DR SMART; SM00369; LRR\_TYP.

CC KW Glycoprotein, Extracellular matrix; proteoglycan; Repeat;

CC KW Leucine-rich repeat; Signal.

CC SIGNAL 1 19 POTENTIAL.

CC PROFP 20 37 BY SIMILARITY.

FT	CHAIN	369	BIGLYCAN.
FT	DOMAIN	64	CYS-RICH.
FT	REPEAT	83	LRR-T 1.
FT	REPEAT	104	LRR-T 127
FT	REPEAT	128	LRR-T 2.
FT	REPEAT	152	LRR-T 2.
FT	REPEAT	173	LRR-T 3.
FT	REPEAT	197	LRR-T 4.
FT	REPEAT	222	LRR-T 3.
FT	REPEAT	243	LRR-T 5.
FT	REPEAT	266	LRR-T 6.
FT	REPEAT	290	LRR-S 4.
FT	REPEAT	291	LRR-S 7.
FT	REPEAT	314	LRR-T 8.
FT	REPEAT	344	O-LINKED (GLYCOSAMINOGLYCAN) (BY SIMILARITY).
FT	CARBOHYD	42	O-LINKED (GLYCOSAMINOGLYCAN) (BY SIMILARITY).
FT	CARBOHYD	48	O-LINKED (GLYCOSAMINOGLYCAN) (BY SIMILARITY).
FT	CARBOHYD	271	N-LINKED (GLCNAC. . . ) (POTENTIAL).
FT	CARBOHYD	312	N-LINKED (GLCNAC. . . ) (POTENTIAL).
FT	DISULFID	64	BY SIMILARITY.
FT	DISULFID	77	BY SIMILARITY.
FT	DISULFID	322	BY SIMILARITY.
FT	DISULFID	369 AA:	41523 MW; 048C8/C233909EE6 CRC64;
SQ	SEQUENCE	369 AA:	
Query Match	51.68;	Score 1027.5;	Length 369;
Best Local Similarity	51.98;	Pred. No. 4.5e-4;	
Matches 196; Conservative	67;	Mismatches 96;	Gaps 3;
QY 4 YVLLFLALCSAKPFPSHIALKNNMLKMDMTDDDD--DDDDDDEDNSLFPTR 59			
Db 5 WLLAALLAASQALPFEQ --- --- --- --- KAFMDFTLDDGLEMLNDEASCAETSGIPDL 53			
QY 60 EPRSHFFPDLFPMCPFGCQCYSRVVHCSDGLTSVPTNIPDFTRMLDQNLKIKEN 119			
Db 54 DS---LPPTYSAMCPFGCQHRSVQCSDSLGEKAVPK1SPDTLLDQNLDSLRKD 109			
QY 120 DFKGILTYGLLUNNKLTKHPKAFELTTKURRLYLISHQNQLEIPNLPKSLAURIE 179			
Db 110 DFKGILQHLYALVLYNNKLSKTHEKAFSPLRKQLQKLYISKNHLVEIPNLPSSLVELRIHD 169			
QY 180 NKVKTIOKDTEFGKNNALIVLEM SANPLDNGTEPAGAFGVTVFHIRLAEEKLTSPKGLP 239			
Db 170 NRKVYPKGESGLRNMCINEMGNCLPQGADGKLNLVRLSEAKLTGPKDLP 229			
QY 240 PTLLPLHLDQWKISTVLELDFKRYKELQGLGNKTDIENGSLANIPREIILENNK 299			
Db 230 ETLINEHLDNKQAEFDLLRYSKLRLGLHQIRMINGSLSPTRLEHLDNNK 289			
QY 300 LKKPSGLPLPEKLYQIILFLHSNSTARYGVNDFCPTVPKMKLKSYASISLPNNPNPVKYWEMQ 359			
Db 290 LSVPAGLPLQVVLQVVLITNNTRKVGVNDFCPVGFVKRAYINGISLFLNNPNPVWEVQ 349			

RESULT 9

ID	PGS1_CANFA	STANDARD;	PRT;	369 AA.
AC 002678;				
DT 15-JUL-1998 (Rel. 36, Created)				
DT 15-JUL-1998 (Rel. 36, Last sequence update)				
DT 15-JUN-2002 (Rel. 41, Last annotation update)				
DE Biglycan precursor (Bone/cartilage proteoglycan I) (PG-S1).				
GN Canis familiaris (Dog).				
OS Euksarota; Melacoda; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC Mammalia; Butheria; Carnivora; Fissipedia; Canidae; Canis.				
NCB1 TAXID:9615;				
RN [1]				
RP SEQUENCE FROM N.A.				
RA Giant T.T.;				

RL	Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.	QY	180 NKVKKIQKDTFKGMNALHVLLENSANPLDNNGTEPGAFEGTVFHIRIAEAKLTSPKGKPGLP 239
CC	-!- SUBUNIT: May be involved in collagen fiber assembly (By similarity).	Db	170 NNRKVPKGVFSLRNNNCTENGNPYLEPGAFGLKLNYLRISEAKLTGIPKDLP 229
CC	-!- SUBCELLULAR LOCATION: Extracellular matrix (By similarity).	Qy	240 PTLLIELHDYKNLKSTVLEDKRYELQRQLGNNKTTDINGSTANPRYRETHLENNK 299
CC	-!- PM: The two attached glycosaminoglycan chains can be either chondroitin sulfate or dermatan sulfate (By similarity).	Db	230 ETLNELHDHDKTQATELDLRLSKYRLGHNQTRMENGSLSFLPTRELHDLNK 289
CC	-!- SIMILARITY: BELONGS TO THE SMALL LEUCINE-RICH PROTEOGLYCANS (SLRPs) FAMILY, CLASS I SUBFAMILY.	QY	300 LKKPSGLPELKYLQLTFLHNSNSTARVGVNDFCPTVPKMKSLYSALISLNNPQVWEMQ 359
CC	-!- SIMILARITY: CONTAINS 12 LEUCINE-RICH REPEATS (LRR).	Db	290 LSVPVSGLPDILQLQQVYLPNNNTKVGVNDFCPVGFVKAAYNGISLNNPQVWEMQ 349
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to license@isb-sib.ch).	QY	360 PATERCIVLSRMVQLQNF 377
CC	EMBL: U83140; AAB51244.1;	Db	350 PATERCIVTDRLAIQFGNY 367
DR	InterPro: IPR01611; LRR.	RESULT 10	
DR	InterPro: IPR000372; LRR_Nterm.	ID	PGS1_BOVIN
DR	InterPro: IPR003592; LRR_out.	ID	P21809; P79259;
DR	InterPro: IPR013591; LRR_typ.	AC	01-MAY-1991 (Rel. 18, Created)
DR	Pfam: PF00560; LRR_7.	DT	01-MAY-1998 (Rel. 36, Last sequence update)
DR	Pfam: PF01462; LRRNT; 1.	DT	15-JUN-2002 (Rel. 41, Last annotation update)
DR	SMART: SM00013; LRR; 3.	DE	Biglycan Precursor (Bone/cartilage proteoglycan I) (PG-S1) (Leucine-rich PG I).
DR	SMART: SM00369; LRR_TYP; 1.	GN	BGN
KW	Glycoprotein; Extracellular matrix; Proteoglycan; Repeat;	OS	Bos_taurus (Bovine).
KW	Leucine-rich repeat; Signal.	OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
FT	SIGNAL 1 19	OC	Namibia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
FT	POTENTIAL 20	OC	Bovidae; Bovinae; Bos.
FT	PROPEP 37	OX	NCBI_TaxID=9913;
FT	CHAIN 38	RN	[1]
FT	DOMAIN 64	RP	SEQUENCE FROM N.A.
FT	REPEAT 83	RC	TISSUE=Aorta;
FT	REPEAT 103	RX	MEDLINE=96113563; PubMed=8673009;
FT	REPEAT 104	XU	J.H., Radhakrishnamurthy B., Srinivasan S.R., Berenson G.S.;
FT	REPEAT 128	RA	"Primary structure of bovine aorta biglycan core protein deduced from cloned cDNA.";
FT	REPEAT 152	RT	RT cloned cDNA.";
FT	REPEAT 172	RL	RT Biochem. Mol. Biol. Int. 37:263-272(1995).
FT	REPEAT 173	RN	[2]
FT	REPEAT 197	RN	SEQUENCE OF 38-369.
FT	REPEAT 222	RP	RT TISSUE=Cartilage;
FT	REPEAT 224	RC	RX MEDLINE=8225532; PubMed=2656687;
FT	REPEAT 243	RX	RA Neame P.J., Choi H.U., Rosenberg L.C.;
FT	REPEAT 267	RA	RT "The primary structure of the core protein of the small, leucine-rich proteoglycan (PG I) from bovine articular cartilage.";
FT	REPEAT 290	RT	RT proteoglycan (PG I) from bovine articular cartilage.";
FT	REPEAT 291	RL	RL Biochim. Mol. Biol. Chem. 264:8653-8661(1989).
FT	REPEAT 314	RN	[3]
FT	REPEAT 344	RN	SEQUENCE OF 38-63.
FT	CARBOHYD 42	RP	RX TISSUE=Cartilage;
FT	CARBOHYD 42	RX	RA MEDLINE=89123388; PubMed=2914936;
FT	CARBOHYD 48	RA	RA Cho J.H.U., Johnson T.L., Pai S., Wang L.H., Rosenberg L.C.,
FT	CARBOHYD 48	RA	RA Neame P.J.;
FT	CARBOHYD 271	RT	RT "Characterization of the dermatan sulfate proteoglycans, DS-PGI and DS-PGII, from bovine articular cartilage and skin isolated by octyl-
FT	CARBOHYD 312	RT	RT sepharose chromatography";
FT	CARBOHYD 64	RT	RT J. Biol. Chem. 264:2875-2884(1989).
FT	DISULFID 322	RT	RT [4]
FT	DISULFID 322	RN	INTERACTIONS WITH MFAP2 AND ELN.
FT	SEQUENCE 369 AA;	RN	RA Reinboth B., Hanssen E., Cleary E.G., Gibson M.A.;
FT	SEQUENCE 369 AA;	RN	RT Molecular interactions of biglycan and decorin with elastic fiber components: biglycan forms a ternary complex with tropoelastin and microfibril-associated glycoprotein 1.;"
FT	Best Local Similarity 51.38%; Score 1021.5; DB 1; Length 369;	RA	RA J. Biol. Chem. 277:3950-3957(2002).
FT	Matches 195; Conservative 68; Mismatches 96; Indels 19; Gaps 3;	RA	CC -!- FUNCTION: May be involved in collagen fiber assembly.
FT	4 YLLFLIALCSNKPPSPSHIALKNNMLKDMDTDDDDDEDDNSLFPTR 59	CC -!- SUBUNIT: Forms a ternary complex with MFAP2 and ELN.	
DB	5 WIYASULLALSQALPFEQ-----KFWDFTLUDGLPLMLNBEASSAETISGVPDL 53	CC -!- SUBCELLULAR LOCATION: Extracellular matrix.	
Oy	60 EPRSHFFPFDLPMCPGCGOYCERYVHCSLDLIGLTSVPTNIPFDTRMLDQNKKIKEKEN 119	CC -!- TISSUE SPECIFICITY: Found in several connective tissues, especially in articular cartilages.	
DB	54 DALPTVHS---AMCPFGCHCHLRVQCSDGLKAVPKKEISPDMLLDQNNDISELRAD 109	CC -!- PTM: The two attached glycosaminoglycan chains can be either	
Oy	120 DFKGKTSLYGILLNNNNKLTKHKAFLTIKRLRLYLSHNQLETPLNPKSAELRITH 179	CC -!- SUBCELLULAR LOCATION: Articular cartilage.	
DB	110 DFKGHLHLYALVNNNNKISKHEKASPLRKQLEYISKNHVEPPNLPSSLVELRHD 169	CC -!- TISSUE SPECIFICITY: Found in several connective tissues, especially in articular cartilages.	

chondroitin sulfate or dermatan sulfate.

CC -!- SIMILARITY: BELONGS TO THE SMALL LEUCINE-RICH PROTEOGLYCANS

CC (SLRPS) FAMILY. CLASS I SUBFAMILY: LEUCINE-RICH REPEATS (LRR).

CC -!- SIMILARITY: CONTAINS 12 LEUCINE-RICH REPEATS (LRR).

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [License@isb-sib.ch](mailto:License@isb-sib.ch)).

CC EMBL: S82652; AAB46746; 2; -.

DR PIR: A33701; A33701; .

DR InterPro: IPR001611; LRR.

DR InterPro: IPR000372; LRR\_Nterm.

DR InterPro: IPR003592; LRR\_out.

DR InterPro: IPR003591; LRR\_typ.

DR Pfam: PF00560; LRR; 7.

DR Pfam: PE01462; LRRNT; 1.

SMART: SM00370; LRR; 3.

SMART: SM00013; LRRNT; 1.

SMART: SM00369; LRR\_TYP; 1.

GW Glycoprotein; Extracellular matrix; Proteoglycan; Signal; Repeat; Leucine-rich repeat.

FT SIGNAL 1 19

FT PROPEP 20 37

FT - CHAIN 38 369

FT DOMAIN 64 77

FT REPEAT 83 103

FT REPEAT 104 127

FT REPEAT 128 151

FT REPEAT 152 172

FT REPEAT 173 196

FT REPEAT 197 221

FT REPEAT 222 242

FT REPEAT 243 266

FT REPEAT 267 290

FT REPEAT 291 313

FT REPEAT 314 343

FT REPEAT 344 369

FT CARBOHYD 42 42

FT CARBOHYD 48 48

FT CARBOHYD 271 271

FT DISULFID 64 77

FT DISULFID 322 355

FT CONFLICT 152 152

FT CONFLICT 188 188

FT CONFLICT 354 354

FT CONFLICT 368 368

FT SEQUENCE 369 AA: 41509 MW: 45330FFFB01B8872; CRC64:

Query Match 51.28% Score 1019.5; DB 1; Length 369;

Best Local Similarity 51.98% Pred. No. 1.6e-63; Gaps 3;

Matches 195; Conservative 65; Mismatches 97; Indels 19; Gaps 3;

CC chondroitin sulfate or dermatan sulfate.

CC -!- SIMILARITY: BELONGS TO THE SMALL LEUCINE-RICH PROTEOGLYCANS

CC (SLRPS) FAMILY. CLASS I SUBFAMILY: LEUCINE-RICH REPEATS (LRR).

CC -!- SIMILARITY: CONTAINS 12 LEUCINE-RICH REPEATS (LRR).

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [License@isb-sib.ch](mailto:License@isb-sib.ch)).

CC EMBL: AAB46746; 2; -.

DR InterPro: IPR001611; LRR.

DR InterPro: IPR000372; LRR\_Nterm.

DR InterPro: IPR003592; LRR\_out.

DR InterPro: IPR003591; LRR\_TYP.

DR Pfam: PF00560; LRRNT; 1.

SMART: SM00370; LRR; 3.

SMART: SM00013; LRRNT; 1.

SMART: SM00369; LRR\_TYP; 7.

KW Glycoprotein; Extracellular matrix; Proteoglycan; Repeat; Leucine-rich repeat; Signal.

FT SIGNAL 1 15 POTENTIAL.

FT PROPEP 16 29 BY SIMILARITY.

FT CHAIN 30 356 DECORIN.

FT DOMAIN 51 64 CYS-RICH.

Qy 6 LLLFLAAGCSAKPFSPSHIALNMMLKMDETDDDD--DDDDDDDDNSLPPTREP 61

Dr 7 LAALLAISQALPPEQ----KAFWDFTLDGLPMLNDEASGAETTSGIPDLS 55

Qy 62 RSHFFPFDFLPMCPGPGCOCYSSRVVHCSDLGLTSVPTNIPFDRLQNNNIKEIKENDF 121

Db 56 ---LPPYIAMSMPFGCHLRVQCSDLKAVPREISPDTLQLQNNDISELKDDF 111

Qy 122 KGLTSLYGLLNNNNKLTKHPKAFLTTKKLRLYLPNHSNOLSEPLNPKSLAELTHENK 181

Db 112 KGLQHLYALVLYNNKISKIHEKAFLSPRLKQLYISKNHCLTIPPNLQSSVELRHNDR 171

Qy 182 VKKIQLQDFPKGMNALHLEMANSPLDNGIEGAFEGVTYFHRIIAEAKLTSVPLPPT 241

Db 172 IRKVKPGVFSGLRNMMNCIEMGGCNPLENSGEFGAFDGLKLNLYRISAEAKLGGPKDLPET 231





DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DECORIN precursor (Bone proteoglycan II) (PG-S2) (PG40).  
 DE DCN.  
 RN Homo sapiens (Human).  
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OC NCBI\_TaxID=9606;  
 [1] RN  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=87017013; PubMed=3484330;  
 KR Krusius T.; Ruoslahti E.;  
 RA "Primary structure of an extracellular matrix proteoglycan core  
 RT protein deduced from cloned cDNA.";  
 RT Proc. Natl. Acad. Sci. U.S.A. 83:7683-7687(1986).  
 [2] RN  
 RP SEQUENCE FROM N.A.  
 PC TISSUE=LUNG;  
 MEDLINE=93162643; PubMed=8432527;  
 Vetter U.; Vogel W.; Just W.; Young M.F.; Fisher L.W.;  
 RT "Human decorin gene: intron-exon junctions and chromosomal  
 localization";  
 RT Genomics 15:161-168(1993).  
 [3] RN  
 RN SEQUENCE OF 1-70 FROM N.A.  
 RX MEDLINE=93162642; PubMed=8432526;  
 RA Danielson K.G.; Fazio A.; Cohen I.R.; Cannizzaro L.; Iozzo R.V.;  
 CS-Szabo G.; Glant T.T.;  
 RA "The human decorin gene: intron-exon organization, discovery of two  
 RT alternatively spliced exons in the 5' untranslated region, and  
 RT mapping of the gene to chromosome 12q23.";  
 RT Genomics 15:146-160(1993).  
 [4] RN  
 RN SEQUENCE FROM N.A. (ISOFORMS A; B; C; D AND E).  
 RX CS-Szabo G.; Glant T.T.;  
 RA "Alternative splicing of human decorin.";  
 RT Submitted (MAR1999) to the EMBL/GenBank/DBJ databases.  
 [5] RN  
 RN SEQUENCE FROM N.A.  
 RX Strausberg R.;  
 RA Submitted (MAR2001) to the EMBL/GenBank/DBJ databases.  
 [6] RN  
 RN SEQUENCE OF 31-50.  
 RX MEDLINE=90073579; PubMed=22530169;  
 RA Roughley P.J.; White R.J.;  
 RT "Dermatan sulphate proteoglycans of human articular cartilage. The  
 properties of dermatan sulphate proteoglycans I and II";  
 Biochem. J. 262:823-827(1989).  
 [7] RN  
 RN SEQUENCE OF 31-49.  
 RX MEDLINE=87250639; PubMed=3597437;  
 RA Fisher L.W.; Hawkins G.R.; Tuross N.; Termino J.D.;  
 RT "Purification and partial characterization of small proteoglycans  
 and II, bone sialoproteins I and II, and osteonectin from the mandibular  
 compartment of developing human bone.";  
 RT J. Biol. Chem. 262:9702-9708(1987).  
 CC -!- FUNCTION: May affect the rate of fibrils formation.  
 CC -!- SUBUNIT: Binds to type I and type II collagen, to fibronectin  
 CC TGF-beta. Forms a ternary complex with MFAP2 and ELN (By  
 CC similarity).  
 CC -!- SUBCELLULAR LOCATION: Extracellular matrix.  
 CC -!- ALTERNATIVE PRODUCTS: 5 isoforms: A (shown here), B, C, D and  
 CC are produced by alternative splicing.  
 CC -!- PTM: The attached glycosaminoglycan chain can be either  
 CC chondroitin sulfate or dermatan sulfate depending upon the site  
 CC of origin.  
 CC -!- SIMILARITY: BELONGS TO THE SMALL LEUCINE-RICH PROTEOGLYCANS  
 CC (SURPS) FAMILY. CLASS I SUBFAMILY.  
 CC -!- SIMILARITY: CONTAINS 12 LEUCINE-RICH REPEATS (LRR).

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/>) or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).

---

DR EMBL; M14219; AAB00774.1; -.

DR EMBL; L01131; AAA52301.1; ALT SEQ.

DR EMBL; L01125; AAA52301.1; JOINED.

DR EMBL; L01126; AAA52301.1; JOINED.

DR EMBL; L01127; AAA52301.1; JOINED.

DR EMBL; L01129; AAA52301.1; JOINED.

DR EMBL; L01130; AAA52301.1; JOINED.

DR EMBL; N98262; AAB60901.1; -.

DR EMBL; AF138301; AAF61437.1; -.

DR EMBL; AF138301; AAF61437.1; -.

DR EMBL; AF138302; AAD44714.1; -.

DR EMBL; AF138303; AAF61438.1; -.

DR EMBL; AF138304; AAD44715.1; -.

DR PIR; BC005322; AAH05322.1; -.

DR PIR; A26476; NBNUC8.

DR PIR; S05640; S05640.

DR PIR; B28457; B28457.

DR PIR; A45016; A45016.

DR Genew; HGNC:2705; DCN.

DR MIM; 125255;

DR InterPro; IPR001611; LRR.

DR InterPro; IPR000372; LRR\_Nterm.

DR InterPro; IPR003592; LRR\_out.

DR InterPro; IPR003591; LRR\_typ.

DR Pfam; PF00560; LRR; 7.

DR Pfam; PF01462; LRRN7; 1.

DR SMART; SM00013; LRR; 3.

DR SMART; SM00013; LRRN7; 1.

DR SMART; SM00369; LRR\_TYP; 1.

**KW** Glycoprotein; Extracellular matrix; Proteoglycan; Repeat; Leucine-rich repeat; Signal; Alternative splicing; Polymorphism.

FT SIGNAL; 1 16 POTENTIAL.

FT PROPEP . 17 30

FT CHAIN . 31 359 DECORIN.

FT DOMAIN . 54 67 CYS-RICH.

FT REPEAT . 73 93 LRR-S 1.

FT REPEAT . 94 117 LRR-T 1.

FT REPEAT . 118 141 LRR-T 2.

FT REPEAT . 142 162 LRR-S 2.

FT REPEAT . 163 186 LRR-T 3.

FT REPEAT . 187 212 LRR-T 4.

FT REPEAT . 213 233 LRR-S 3.

FT REPEAT . 234 257 LRR-T 5.

FT REPEAT . 258 281 LRR-T 6.

FT REPEAT . 282 304 LRR-S 4.

FT REPEAT . 305 334 LRR-T 7.

FT REPEAT . 335 359 LRR-T 8.

FT DISULFID 313 54 BY SIMILARITY.

FT CARBOHYD 313 67 O-LINKED GLYCOSAMINOGLYCAN.

FT CARBOHYD 34 346 BY SIMILARITY.

FT CARBOHYD 211 211 N-LINKED GLCNAC (POTENTIAL).

FT CARBOHYD 262 262 N-LINKED GLCNAC (POTENTIAL).

FT CARBOHYD 303 303 N-LINKED GLCNAC (POTENTIAL).

FT VARSPLIC 71 179 MISSING (IN ISOFORM B).

FT VARSPLIC 73 219 MISSING (IN ISOFORM C).

FT VARSPLIC 109 295 MISSING (IN ISOFORM D).

FT VARSPLIC 72 75 LDVW > CLPS (IN ISOFORM E).

FT VARSPLIC 76 359 MISSING (IN ISOFORM E).

FT VARIANT 273 273 E > Q (IN DBSNP:1803344).

FT CONFLICT 37 37 G > A (IN REF. 7).

FT CONFLICT 45 45 D > P (IN REF. 7).

SEQUENCE 359 AA; 39746 MW; FF511E871ALA52DD CRC64;

Query Match 48.4%; Score 963.5; DB 1; Length 359;

Best Local Similarity 50.4%; Pred. No. 1..le-59;

Matches 191; Conservative 65; Mismatches 100; Indels 23; Gaps 4;

1 MKYVLLFLALCS-AKPFSPSHIALKNMKLDMETDDDDDDDDNSLELFPTR 59

Db	1	MKATILLLAQVSWAGPF--QQRGLEDFMELDEASGIGPEVPPDRD-----	45	DR	InterPro; IPR01611; LRR_Nterm.
Qy	60	EPRSSHFFPFDLFPMCPFGCOCYSRVHCSDFLGLRSVPTNPFTRMLDLQNNKIKEIKEN	119	DR	InterPro; IPR00372; LRR_Out.
Db	46	--FEPSSGPVCFRCOCHLRVQCSDFLGLDKVPDKLPPDTLQLQNNKITEIKDG	99	DR	InterPro; IPR03591; LRR_Typ.
Qy	120	DFKGLTSLYGLILNNNNKLTKTHPKAFLTTKLLRNOLSHNOLSTPLNPKSLAELRTH	179	DR	Pfam; PF00560; LRR; 1.
Db	100	DFKNKLNLHALLVNNKSKVSGANTPLVLERLYLSKQNOLKEPLMPKTQELRAHE	159	DR	SMART; SMART000370; LRR; 2.
Qy	180	NKVVKQKDIFPKGMNALHVLEMSANPDLNDNGTEPCAGEFY-TVPHIRIAEKATSVPKG	238	DR	SMART; SMART000369; LRR_Typ; 2.
Db	160	NEITKVKVTNGLNQMVIEGTPLKSSGTEAFOGMKKLSSYTRADNTNTSIPQGL	219	KW	Glycoprotein; Extracellular matrix; Proteoglycan; Repeat;
Qy	239	PPTLLELHDYNIKISTVELEDKRYKTELQLRLGQNKTDIENGSLANTPVRTEIHENN	298	KW	Leucine-rich repeat; Signal_Potential.
Db	220	PSSLELHDGDKISRVDAASLKGQLNLGSNSISAHDNGSLANTPHREHLDDNN	279	FT	PROPEP
Qy	299	KLUKPSGLPELKYLOLTIIFLHNINSTARGVYNDLISLNFNPVKYWM	358	FT	DOMAIN
Db	280	KLTRVPGLAGBRYIIVYLNINNSVVGSSDFCPGHNTKKASYSGVSLFSNFVQYWEI	339	FT	REPEAT
Qy	359	QPATERCVLSRMVQLGNF	377	FT	REPEAT
Db	340	QPSTERCVYVRSQAIQIGNY	358	FT	REPEAT
<hr/>					
RESULT 15					
ID	PGS2_SHEEP	STANDARD;	PRT;	360 AA.	FT
AC	Q9TE2;				CARBOHYD
DT	15-JUN-2002	(Rel. 41, Created)			FT
DT	15-JUN-2002	(Rel. 41, Last sequence update)			CARBOHYD
DT	15-JUN-2002	(Rel. 41, Last annotation update)			FT
DE	Decorin precursor (Bone proteoglycan II) (PG-S2) (FG40).				SQ
GN	DCN.				SEQUENCE
OS	Ovis aries (Sheep).				
OC	Mammalia; Chordata; Craniata; Vertebrata; Ruminantia; Pecora; Bovidae;				
OC	Eutheria; Cetartiodactyla; Caprinae; Ovis.				
OC	NCBI TaxID=9940;				
OX	[1]				
RN	RP				
RC	SEQUENCE FROM N.A.				
RX	MEDLINE-20112292; PubMed=110644528;				
RA	Wu W.X., Zhang Q., Unno N., Derkis J.B., Nathanielsz P.W.;				
RT	"Characterization of decorin mRNA in pregnant intrauterine tissues of the ewe and regulation by steroids."				
RL	J. Physiol. 278:C199-C206(2000).				
-!- FUNCTION:	May affect the rate of fibrils formation (By similarity).				
-!- SUBUNIT:	Binds to type I and type II collagen, to fibronectin and TGF- $\beta$ 1. Forms a ternary complex with MFAP2 and ELN (By similarity).				
CC	SUBCELLULAR LOCATION: Extracellular matrix (By similarity).				
CC	-!- PTM: The attached glycosaminoglycan chain can be either chondroitin sulfate or dermatan sulfate depending upon the tissue of origin (By similarity).				
CC	-!- SIMILARITY: BELONGS TO THE SMALL LEUCINE-RICH PROTEOGLYCANS (SLRPs) FAMILY. CLASS I SUBFAMILY.				
CC	-!- SIMILARITY: CONTAINS 12 LEUCINE-RICH REPEATS (LRR).				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license/agreement (see <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to license@isb-sib.ch).				
CC	-----				
EMBL; AAF125041; AAF00585.1; -;					
HSSP; P09661; 1AQ9.					
DR	InterPro; IPR01611; LRR_Nterm.				
DR	InterPro; IPR00372; LRR_Out.				
DR	InterPro; IPR03591; LRR_Typ.				
DR	Pfam; PF00560; LRR; 9.				
DR	SMART; SMART000370; LRR; 2.				
DR	SMART; SMART000369; LRR_Typ; 2.				
FT	Glycoprotein; Extracellular matrix; Proteoglycan; Repeat;				
FT	Leucine-rich repeat; Signal_Potential.				
FT	BY SIMILARITY.				
FT	PROPEP				
FT	DECORIN.				
FT	CYS-RICH.				
FT	REPEAT				
FT	LRR_S; 1.				
FT	REPEAT				
FT	94				
FT	118				
FT	142				
FT	163				
FT	187				
FT	213				
FT	234				
FT	258				
FT	282				
FT	305				
FT	335				
FT	360				
FT	DISULFID				
FT	55				
FT	68				
FT	BY SIMILARITY.				
FT	O-LINKED GLYCOSAMINOGLYCAN (BY SIMILARITY).				
FT	N-LINKED (GLCNAC. . . ) (POTENTIAL).				
FT	N-LINKED (GLCNAC. . . ) (POTENTIAL).				
FT	N-LINKED (GLCNAC. . . ) (POTENTIAL).				
FT	N-LINKED (GLCNAC. . . ) (POTENTIAL).				
FT	0095DODFDAB88624 CRC64;				
FT	39972 MW;				
FT	3960 AA;				
Query Match	47.5%	Score 947;	DB 1;	Length 360;	
Best Local Similarity	49.78;	Pred. No. 1.6e-58;			
Matches 190; Conservative 67; Mismatches 97; Indexes 28; Gaps 7;					
Qy	1 MKEYVLLPLALCS-AKPFSPSPHTALKMMKLDMEDDDDDDDDDDDNSELPTR 59				
Db	1 MKATIIFFLVQAVQSMWGPF--QOKGLFEFML-----EDEASGIGP-				
Qy	60 EPRSSHEFFPDLPFM--DPLGCTYSSRVSRLVQSDLGLTSVPTNIPFDTMLDLQNNKIKEI 116				
Db	39 EERFIEVP-ELEPGPVCERQCHLRVQCSDFGLERVKPDIPDTALDLQNNKIKEI 97				
Qy	117 KENDFKGLTSLYGLILNNNNKLTKHPKAFLTTKLRRLYSLHNOLOSEPLNPKSLAELR 176				
Db	98 KDGFDFKLNKLNHTLILNNKISKSPGAFLYLERLYLSKNOKELEPKTQLER 157				
Qy	177 THENVKKIKDQKDFKGMLAHVLEMSANLDNNGIEPGFEGV-TVFHTRIAEAKLTSVP 235				
Db	158 VHENEITKVKSVFNGLNQMVVEGTPLNPKSKTENGAFQGMKLSYTRIAQTNTTP 217				
Qy	236 KGLPPLIELHDYNNKITSVTELEDKRYKELRQUGLGNNKTTDLENGSLANIPIVRETHL 295				
Db	218 QGLPSSLTELHLDGNNKITYKVDAASLKGGLNNAKGLSFNSISAYDNGSLANTPHRELHL 277				
Qy	296 ENNKKKTPSGLPELKYLOLFLHSNSIARVGNDFCPVPKMKSLYSAISLFNPVYK 355				
Db	278 NNKLVYKPGSLADHKYIQVYLYHNNNTSAISGNDFCPGYNTKASVGSVLFSNPVYQ 337				
Qy	356 WEMOPATFRCVLSRMSVQLGNF 377				
Db	338 WEIOPSTERCVYRAVQLGNY 359				
Search completed: May 27, 2003, 16:04:27					
Job time : 16 secs					

GenCore version 5.1.4-p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

protein - protein search, using sw model

on: May 27, 2003, 15:46:26 ; Search time 37 Seconds  
(without alignments)  
1364.919 Million cell updates/sec

file: US-09-944-457-2  
refect score: 1992  
quence: 1 MKEYVLLFLALCSAKPFFS.....PATERCVLSRMSVQLGNFGM 379

oring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

arched: 908470 seqs, 133250620 residues

real number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

st-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

atabase :

A-Geneseq_101002,*	1: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT:*	3: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT:*	5: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT:*	7: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT:*	9: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT:*	11: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT:*	13: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT:*	15: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT:*	17: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT:*	19: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT:*	21: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT:*	23: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

8

Result No.	Score	Match Length	DB ID	Description
1	1992	100.0	379	21 AAB01311 Human PR0241 polyprotein
2	1992	100.0	379	22 AAU2335 Human PR0241 polyprotein
3	1992	100.0	379	23 ABB05437 Human PR0241 proteoglycan
4	1992	100.0	379	23 ABB04831 Human PR0241 protein
5	1990	99.9	379	20 AAY17820 Human PR0241 protein
6	1799	90.3	373	22 AAE05347 Mouse bone/cartilage
7	1799	90.3	373	23 ABB2356 Murine protein isoform
8	1649	82.8	344	22 AAM0351 Human polypeptide
9	1558	78.2	352	22 ABG2569 Novel human dipeptidyl peptidase inhibitor
10	1237	62.1	246	22 AAB05678 Human protein sequence

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	Peptide	label
1	1992	100.0	379	21 AAB01311	Human PRO241 polyp	FT	1.15
2	1992	100.0	379	21 AAU12335	Human PRO241 polyp	FT	129..135
3	1992	100.0	379	23 ABB95437	Human angiogenesis	FT	note= "N-myristoylation site"
4	1992	100.0	379	23 ABB84811	Human PRO241 prote	FT	154..176
5	1990	99.9	379	20 AAY17820	Human PRO241 prote	FT	note= "Leucine zipper pattern
6	1799	90.3	373	22 AAE05347	Mouse bone/cartila	FT	210..216
7	1799	90.3	373	23 ABB72356	Murine protein iso	FT	note= "N-myristoylation site"
8	1649	82.8	344	22 AAM40351	Human polypeptide	FT	214..220
9	1558	78.2	352	22 ABG22569	Novel human diagno	FT	note= "N-myristoylation site"
10	1237	62.1	246	22 AAB95678	Human protein sequ	FT	237..243

FT	Modified-site	/note= "N-myristoylation site"
FT	Modified-site	/note= "N-glycosylation site"
FT	Modified-site	/note= "N-myristoylation site"
XX		
PN	WO200032776-A2.	
PD	08-JUN-2000.	
XX		
PF	01-DEC-1999;	99WO-US28301.
XX		
PR	01-DEC-1998;	98WO-US25108.
PR	16-DEC-1998;	98US-0112450.
PR	22-DEC-1998;	98US-0113296.
XX		
PA	(GETH ) GENENTECH INC.	
XX		
PI	Baker KP, Botstein D, Eaton DL, Ferrara N, Filvaroff E;	
PI	Gerritsen ME, Goddard A, Godowski PJ, Grimaldi CJ, Gunney AL;	
PI	Hillman KJ, Kljavin IJ, Napier MA, Roy MA, Tomas D, Wood WI;	
XX		
WPI;	2000-412324/35.	
XX		
PR	New human nucleic acids encoding secreted and transmembrane polypeptides, designated as PRO polypeptides, useful as pharmaceutical and diagnostic agents	
PR	Claim 12; Fig 2; 187pp; English.	
PS		
XX		
CC	New human nucleic acids encoding secreted and transmembrane polypeptides which are designated as PRO polypeptides, are described including as membrane-bound proteins have various industrial applications, including as pharmaceutical and diagnostic agents. The membrane-bound proteins can also be employed for screening of potential peptide or small molecule inhibitors of the relevant receptor/ligand interaction. Anti-PRO antibodies are useful for the affinity purification of PRO from recombinant cell culture or natural sources.	
CC		
XX		
Sequence	379 AA:	
Query Match	100..08;	Score 1992; DB 21; Length 379;
Best Local Similarity	100..08;	Pred. No. 4e-177;
Matches	379;	Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy	1	MREYVLLFLAACSARKPFFSPSHIAALKNMILKDMEDTDDDDDDDDDENSLFPTR 60
Db	1	MREYVLLFLAACSARKPFFSPSHIAALKNMILKDMEDTDDDDDDDDDENSLFPTR 60
61	PRSHFEPFDLPMCPGCGCYSRVHCSIDGLTSVPTNIPFDTRMLDQNKKKEITKEND 120	
61	PRSHFEPFDLPMCPGCGCYSRVHCSIDGLTSVPTNIPFDTRMLDQNKKKEITKEND 120	
Oy	121	FKGLTSLYGLTANNKLNKTLKKAFLTTLKRLYLSHNOSETPLNLPKSLAEIIRHEN 180
Db	121	FKGLTSLYGLTANNKLNKTLKKAFLTTLKRLYLSHNOSETPLNLPKSLAEIIRHEN 180
Oy	181	KVKRKIQDTFGKMNALHVLEMSANLDNGTEPGAFEGTVFHRIAEEAKLTSVPKGLEPP 240
Db	181	KVKRKIQDTFGKMNALHVLEMSANLDNGTEPGAFEGTVFHRIAEEAKLTSVPKGLEPP 240
Oy	241	TLEELHDYDKNSTVLEDFPKRYKELORUGLGNNTIDENGSLANIPRVEITHLENNKL 300
Db	241	TLEELHDYDKNSTVLEDFPKRYKELORUGLGNNTIDENGSLANIPRVEITHLENNKL 300
Oy	301	KKIPSGLPELYLQLQIFLHSNSIARYGVNDFCPTPKMKSLYSAYSISLNNPQKYWEMQP 360
Db	301	KKIPSGLPELYLQLQIFLHSNSIARYGVNDFCPTPKMKSLYSAYSISLNNPQKYWEMQP 360
Oy	361	ATFRCVLSRMSVQLGNFGM 379
Db	361	ATFRCVLSRMSVQLGNFGM 379

CC PRO polypeptides, and to detect the presence of mammalian lung, colon, breast, prostate, rectal, cervical or liver tumours by comparing PRO polypeptide expression in a cell sample to that in a control sample. Some of the 275 sequences are also useful to stimulate the release of tumour necrosis factor-alpha (TNF-alpha) from human blood, the proliferation or differentiation of chondrocytes, the proliferation or gene expression in pericyte cells, the release of proteoglycans from cartilage, the proliferation of inner ear utricular supporting cells or of T-lymphocytes, the release of a cytokine from peripheral blood monocytes (PBMCs), or the proliferation of endothelial cells. Some of the PRO polypeptides may modulate glucose or free fatty acid uptake by skeletal muscle cells or by adipocytes; or inhibit binding of A-peptide to factor VIIA. The PRO polypeptides can be used in assays to identify molecules involved in binding interactions. The polynucleotides encoding PRO polypeptides can be used to generate probes, antisense RNA/DNA, transgenic or knock out animals and can be used in gene therapy.

## Sequence

379 AA;

Query Match 100.0%; Score 1992;

Best Local Similarity 100.0%; Pred. No. 4e-177;

Length 379;

Matches 379;

Conservative 0;

Mismatches 0;

Indels 0;

Gaps 0;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

CC hypertrophy; trauma; cancer; age-related macular degeneration; atherosclerosis; hypertension; arterial restenosis; rheumatoid arthritis; angina; myocardial infarction; thromboangiitis; lymphangiitis; tumour angiogenesis (such as breast carcinoma and liver carcinoma) and wound healing. The present sequence is a PRO protein of the invention.

XX Sequence 379 AA;

Query Match 100.0%; Score 1692; DB 23; Length 379;

Best Local Similarity 100.0%; Pred. No. 4e-177; Mismatches 0; Indels 0; Gaps 0;

Matches 379; Conservative 0; Nucleotides 0;

CC atherosclerosis; hypertension; arterial restenosis; rheumatoid arthritis;

CC angina; myocardial infarction; thromboangiitis; lymphangiitis; tumour

CC angiogenesis (such as breast carcinoma and liver carcinoma) and wound

CC healing. The present sequence is a PRO protein of the invention.

XX

SQ Sequence 379 AA;

PR 02-AUG-2000; 20000US-222695P.

PR 17-AUG-2000; 20000US-064357.

PR 23-AUG-2000; 20000WO-US23522.

PR 24-AUG-2000; 20000WO-US23438.

PR 07-SEP-2000; 20000US-23098P.

PR 18-SEP-2000; 20000US-0664510.

PR 18-SEP-2000; 20000US-0665350.

PR 24-OCT-2000; 20000US-242922P.

PR 08-NOV-2000; 20000US-0709238.

PR 08-NOV-2000; 20000WO-US30932.

PR 10-NOV-2000; 20000WO-US30873.

PR 01-DEC-2000; 20000WO-US32678.

PR 20-DEC-2000; 20000US-0747259.

PR 20-DEC-2000; 20000WO-US34956.

PR 22-JAN-2001; 20001US-0767609.

PR 28-FEB-2001; 20001US-0796499.

PR 05-APR-2001; 20001US-0828366.

PR 10-MAY-2001; 20001US-0854208.

PR 01-MAR-2001; 20001US-0806666.

PR 09-MAR-2001; 20001US-0802706.

PR 14-MAR-2001; 20001US-0808689.

PR 22-MAR-2001; 20001US-0816744.

PR 28-MAR-2001; 20001US-0823661.

PR 05-APR-2001; 20001US-0828366.

PR 10-MAY-2001; 20001US-0854280.

PR 25-MAY-2001; 20001US-0866028.

PR 25-MAY-2001; 20001US-0866034.

PR 25-MAY-2001; 20001US-0817092.

PR 30-MAY-2001; 20001US-0870574.

PR 01-JUN-2001; 20001US-0817443.

PR 01-JUN-2001; 20001WO-US17800.

XX

PA (GETH ) GENENTECH INC.

XX

PI Baker KP, Ferrara N, Gerber H, Gariel M, Goddard A, Godowski PJ, Gurney AL, Hillan KJ, Marsters SA, Pan J, Paoni NF;

PI Stephan JP, Watanabe CK, Williams PM, Wood WI, Ye W;

XX

DR WPI; 2002-090516/12.

DR N-PSDB; ABL88086.

XX

PT One hundred and eighty seven nucleic acids encoding PRO polypeptides,

PT useful in diagnosis and treatment of cardiovascular (e.g. myocardial

PT infarction), endothelial or angiogenic disorders in a mammal -

XX

CC Claim 11; Fig 30; 565PP; English.

XX

CC ABL88072 to ABL88258 encode the PRO proteins given in ABB84817 to

CC ABB85003. The PRO proteins and polynucleotides have cardiotonic, cytostatic,

CC antiangiogenic, hypotensive, vasoactive and antiarteriosclerotic

CC activities, and can be used in gene therapy. The PRO polynucleotides,

CC proteins, agonists and antagonists are useful for treating or diagnosing

CC a cardiovascular, endothelial or angiogenic disorder in a mammal

CC e.g., cardiac hypertrophy, trauma, cancer, age-related macular

CC degeneration, atherosclerosis, hypertension, arterial restenosis,

CC angioid arteritis, angioma, myocardial infarctions, thrombophlebitis,

CC lymphangitis, tumour angiogenesis (such as breast carcinoma and liver

CC carcinoma) and wound healing. The PRO polynucleotides have applications

CC in molecular biology, including use as hybridisation probes, and in

CC chromosome and gene mapping. ABL88259 to ABL88267 represent primers and

CC probes used in the exemplification of the present invention.

XX Sequence 379 AA;

XX

CC Query Match 100.0%; Score 1992; DB 23; Length 379;

CC Best Local Similarity 100.0%; Pred. No. 4e-177;

CC Mismatches 0; Indels 0; Gaps 0;

CC Matches 379; Conservative 0;

CC

DE Human PRO241 protein sequence SEQ ID NO:30.

XX Human PRO241 protein sequence SEQ ID NO:30.

XX

AC ABB84831;

XX 16-MAY-2002 (first entry)

XX

DE Human PRO241 protein sequence SEQ ID NO:30.

XX

Human; angiogenesis; cardiotonic; cytostatic; antiangiogenic; hypotensive;

KW vascular; antiarteriosclerotic; PRO agonist; trauma;

KW gene therapy; cardiovascular disorder; endothelial disorder; cancer;

KW angiogenic disorder; cardiac hypertension; atherosclerosis;

KW age-related macular degeneration; arterial restenosis; angina;

KW rheumatoid arthritis; myocardial infarction; thrombophlebitis;

KW lymphangiitis; tumour angiogenesis; breast carcinoma;

KW wound healing; chromosome mapping; gene mapping.

XX Homo sapiens.

XX

PN WO200200690-A2.

XX 03-JAN-2002.

XX

PF 20-JUN-2001; 2001WO-US19692.

XX

23-JUN-2000; 2000US-213637P.

XX

PR 20-JUL-2000; 2000US-219558P.

XX

25-JUL-2000; 2000US-220664P.

XX

28-JUL-2000; 2000WO-US20710.

XX

PR SHEFPEDLFPMCPFGCQCYSRVHSGDGLTSVPNTPDFTRMLDQNKKIKEKEND 120

XX

PR 1 MKEVVLLFLALCSAKPFFSPSHIALKNNMLKDMEDTDDDDDDNSLFPTR 60

XX

PR

Db	61	PRSHFPFDLFPMPFGCOCYSRVVCSDIGLTSPVNTIPFDTRMLDQNNKIKEIKEND	120
Qy	121	FKGTSLYGLIINNNKLTKHKAFLTTLKLRRLYLSHNOLSEIPLNPKSLAEDRHHN	180
Db	121	FKGTSLYGLIINNNKLTKHKAFLTTLKLRRLYLSHNOLSEIPLNPKSLAEDRHHN	180
Qy	181	KVKKTQDFKGNAHLVLEMSANPLDNGIEPAGFEGTVFHTRIAEAKLTSPKGLLPP	240
Db	181	KVKKTQDFKGNAHLVLEMSANPLDNGIEPAGFEGTVFHTRIAEAKLTSPKGLLPP	240
Qy	241	TLEELHDYNKISTVELEDKFKYKELQRLGNNKNTIDENGSLANIPRVRETHLENNKL	300
Db	241	TLEELHDYNKISTVELEDKFKYKELQRLGNNKNTIDENGSLANIPRVRETHLENNKL	300
Qy	301	KKTPSGLPELKYLQIIFLHSNSTARVGVDNFCPTVKMKKSLYSATISLNNPKRYWENQP	360
Db	301	KKTPSGLPELKYLQIIFLHSNSTARVGVDNFCPTVKMKKSLYSATISLNNPKRYWENQP	360
Qy	361	ATFRCVLSRMSVQLGNFGM	379
Db	361	ATFRCVLSRMSVQLGNFGM	379
		RESULT 5	
		AAV17820	
		AAV17820 standard; Protein; 379 AA.	
		XX	
		12-AUG-1999 (first entry)	
		Human PRO241 protein sequence.	
		Human; PRO protein; tumour necrosis factor family; TNF; cytokine; secreted protein; transmembrane protein; inflammation disorder.	
		Homosapiens.	
		W09928462-A2.	
		XX	
		01-DEC-1998 . 10-JUN-1999.	
		98WO-US25108.	
		XX	
	25-FEB-1998;	98US-0075945.	
	03-DEC-1997;	97US-0067411.	
	11-DEC-1997;	97US-0069278	
	11-DEC-1997;	97US-0069334.	
	11-DEC-1997;	97US-0069435.	
	12-DEC-1997;	97US-0069425.	
	16-DEC-1997;	97US-0069694.	
	16-DEC-1997;	97US-0069696.	
	16-DEC-1997;	97US-0069702.	
	17-DEC-1997;	97US-0069870.	
	17-DEC-1997;	97US-0069873.	
	18-DEC-1997;	97US-0068017.	
	05-JAN-1998;	98US-0070440.	
	09-FEB-1998;	98US-0074086.	
	09-FEB-1998;	98US-0074092.	
	XX	PA (GETH ) GENENTECH INC.	
	XX	Baker KP, Chen J, Goddard A, Gurney AL, Wood WI;	
	XX	Yuan J;	
	XX	WPI; 1999-371118/31.	
	XX	N-PSDB; AA80043.	
	XX	Nucleic acids encoding PRO secreted and transmembrane proteins	
	XX	Claim 12; Fig. 2; 123pp; English.	
	XX	The present invention describes nucleic acids encoding PRO secreted and transmembrane proteins used therapeutically. The PRO proteins have cytosstatic, anti-inflammatory, anti-proliferative and immunosuppressive activity. The proteins and polynucleotides can be used in therapy identification of homologues, raising antibodies and design of probes and primers. They can be used in a range of diseases related to proteins that they have homology with, e.g. a PRO protein having homology to complement proteins may be used in inflammatory responses.	
	XX	Sequence 379 AA;	
	Query Match 99.9%; Score 1990; DB 20; Length 379;		
	Best Local Similarity 99.7%; Pred. No. 6.2e-177;		
	Matches 378; Conservative 1; Mismatches 0; Indels 0; Gaps 0		
	Qy 1 MKEYVLLFLALCSAKPFSPSHIALKNNMILKDMEDTDODDDDDDDDDNSLFPTR 60		
	Db 1 MKEYVLLFLALCSAKPFSPSHIALKNNMILKDMEDTDODDDDDDDDDNSLFPTR 60		
	Qy 61 PRSHFPFDLFPMCPEGCOCYSRVVCSDIGLTSPVNTIPFDTRMLDQNNKIKEIKEND 120		
	Db 61 PRSHFPFDLFPMCPEGCOCYSRVVCSDIGLTSPVNTIPFDTRMLDQNNKIKEIKEND 120		
	Qy 121 FGKLTSLYGLLNNNNKLTKHKAFLITKLRLRLLSHNOLSEIPLNPKSLAELRHTEN 180		
	Db 121 FGKLTSLYGLLNNNNKLTKHKAFLITKLRLRLLSHNOLSEIPLNPKSLAELRHTEN 180		
	Qy 181 KVKKIQKDFFKGMNALHVLEMANSPLDNGTEPGAFEGTVFHTRIAEAKLTSPVKGLLPP 240		
	Db 181 KVKKIQKDFFKGMNALHVLEMANSPLDNGTEPGAFEGTVFHTRIAEAKLTSPVKGLLPP 240		
	Qy 241 TLEELHDYNKISTVELEDKFKYKELQRLGNNKNTIDENGSLANIPRVRETHLENNKL 300		
	Db 241 TLEELHDYNKISTVELEDKFKYKELQRLGNNKNTIDENGSLANIPRVRETHLENNKL 300		
	Qy 301 KKPSGLPELKYLQIIFLHSNSTARVGVDNFCPTVKMKKSLYSATISLNNPKRYWENQP 360		
	Db 301 KKPSGLPELKYLQIIFLHSNSTARVGVDNFCPTVKMKKSLYSATISLNNPKRYWENQP 360		
	Qy 361 ATFCVCLSRMSVQLGNFGM 379		
	Db 361 ATFCVCLSRMSVQLGNFGM 379		
	RESULT 6		
	AAE05347		
	ID AAE05347 standard; Protein; 373 AA.		
	XX		
	XX	AAE05347; AC AAE05347;	
	XX	DT 12-SEP-2001 (first entry)	
	XX	DE Mouse bone/cartilage proteoglycan I (BGN) protein.	
	XX	KW Mouse; cytosatic; antiinflammatory; immuno-regulatory; tissue integrity; wound healing; immune response; vaccine; cancer; asthma; allergy; cell trafficking; BGN; bone/cartilage proteoglycan I protein; biglycan; PG-SI; secreted protein; therapy.	
	XX	Mus sp.	
	XX	PN WO200148192-A1.	
	XX	PD 05-JUL-2001.	
	XX	PF 21-DEC-2000; 2000WO-NZ00256.	
	XX	PR 23-DEC-1999; 99US-0171678.	
	XX	PR 28-NOV-2000; 2000US-0724864.	
	XX	PA (GENE-) GENESIS RES & DEV CORP LTD.	
	XX	PJ Watson JD, Murison JG;	

Page 6

WPI: 2001-425665/45.  
N/PSDB; AAD0114.

Novel isolated polypeptide useful to isolate corresponding interacting proteins or other compounds, to quantitatively determine levels of interacting proteins or other compounds, and as therapeutic target -

Claim 6: Page 76-77; 101pp; English.

The patent discloses novel polynucleotides and their corresponding proteins which play a major role in induction of growth, cell migration and proliferation, cell-cell interaction and the differentiation of tissue-specific cells. These proteins are important in the maintenance of tissue integrity and thus are important in wound healing. They are useful in various assays to determine the biological activity to raise antibodies to isolate corresponding interacting proteins or other compounds, to quantitatively determine levels of interacting proteins or other compounds, and as therapeutic target in a whole range of disease states. Compositions comprising the novel proteins of the invention are useful for treating mammalian disorders. Polynucleotides of the invention are useful in genome and physical mapping, in positional cloning of genes, to tag or identify an organism or its reproductive material (as non-disruptive tags for marking organisms), and for the diagnosis and treatment of mammalian diseases which is the consequence of inappropriate expression of kinase genes. They are useful for promoting immune responses as part of a vaccine or anti-cancer treatment, as target for cancer treatment, as immunoregulatory and anti-inflammatory molecule as diagnostic for specific types of cancer and for development of an anti-cancer treatment, and as a target for antagonists in the treatment of diseases such as asthma and allergy. They are also useful to inhibit or enhance the activity of the soluble molecule that binds proteins of the invention, for tissue and neural regeneration, to promote or block cell trafficking and as anti-inflammatory and/or vaccine adjuvant.

The present sequence is bone/cartilage proteoglycan I (BGN) protein, a secreted protein from mouse. BGN is also known as biglycan or PG-S1.

ID	ABB72356	standard; protein;	373 AA.
XX	AC		
XX	XX	Murine protein isolated from skin c	
XX	DE		
XX	XX	Human: rat; mouse; skin cell; skin	
XX	KW	developmental defect; inflammatory	
XX	KW	immunomodulator; anti-inflammatory	
XX	KW		
XX	OS		
XX	SP		
XX	XX	Mus sp.	
XX	PN	WO200190357-A1.	
XX	XX		
XX	PD	29-NOV-2001.	
XX	XX		
XX	PF	24-MAY-2001; 2001W0-NZ00099.	
XX	XX		
XX	PR	24-MAY-2000; 2000US-206650P.	
XX	PR	25-JUL-2000; 2000US-221232P.	
XX	PA	(GENE-) GENESTS RES & DEV CORP LTD.	
XX	PA		
XX	PJ	watson JD, Strachan L, Sleeman M,	
XX	XX		
XX	DR	WPI; 2002-122020/16.	
XX	DR	N-PSDB; ABL35045.	
XX	PT	New polynucleotides and polypeptides	
PT	PT	isolated from skin cells, useful for	
PT	PT	growth and developmental defects, i	
PT	PT	modulating immune responses -	
PS	PS		
Claim 4:	Page 434-435;	466pp;	English

immunomodulator; anti-inflammatory; cytosolic, neuroprotective.

Mus sp. XX

OS XX

XX WO200190357-A1.

XX PD 29-Nov-2001.

XX PF 24-MAY-2001: 2001WO-NZ00099.

XX PR 24-MAY-2000: 2000US-206650P.

XX PR 25-JUL-2000: 2000US-221222P.

XX PA (GENE-) GENESTS RES & DEV CORP LTD.

XX PT Watson JD, Strachan L, Sleeman M. Onrust R, Murison JG, Kumble KD; WPI; 2002-122020/16.

DR N-PSBB; ABL35005.

XX New polynucleotides and polypeptides encoded by the polynucleotides isolated from skin cells, useful for treating skin wounds, cancers, growth and developmental defects, inflammatory diseases, or for modulating immune responses.

XX PS Claim 4; Page 434-435; 466pp: English.

XX CC The present invention provides the protein and coding sequences of cDNAs isolated from human, murine and rat skin cell libraries. The sequences can be used in the development of therapeutic agents useful in the treatment of skin diseases, including skin wounds, cancer, growth defects, developmental defects and inflammatory diseases. The proteins have important roles in the induction of hair growth, cell proliferation and cell-cell interaction, in maintaining tissue integrity, in wound healing and in modulating immune responses. The present sequence is a polypeptide of the invention.

XX Sequence 373 AA;

Query	Match	Score	DB	Length
Qy	KEYVLLFLIACSAKPFSPFHIALEKNMKLDMEDDDDDDDDDDEINSLFPTRE	90.38;	23;	373;
Db	MKEYVMLLILAVCSAKPFESSHTALKNMKLDMEDDDDDDDDDNDD-----DNSLFPTKE	90.28;	No. 3..9e-159;	
Matches	341; Conservative	17;	Mismatches	14;
			Indels	6;
			Gaps	1;

Qy 1 FKGTLISGLYLINNNKLTIKHPKAELTTKKLRLYLSHNOLSEIPLNPKSLAELRTHEN 180

Db 1 MKEYVMLLILAVCSAKPFESSHTALKNMKLDMEDDDDDDDDDNDD-----DNSLFPTKE 54

Qy 61 PRSHFFPDLDLPPMCPCGOCQYSRVHCSGDLGTSVPTNPFEDTRMLDQNKKIKEIKEND 120

Db 55 PVNPFFPDLDLPTCPGQCQYHSDGILTGSVPNTPFEDTRMVLQDNKKIKEIKEND 114

Qy 121 KVKTKDQKDTHOMNALHVLEMANSPLDNGIEPGAFEGTVHRIEAKLTSVPKGKLP 240

Db 115 FKGLTSLYATLNNKKLTIKHPKTLTTKLRLLYLSHNOLSEIPLNPKSLAELRTHDN 174

Qy 181 KVKTKDQKDTHOMNALHVLEMANSPLDNGIEPGAFEGTVHRIEAKLTSVPKGKLP 234

Db 175 KVKTKDQKDTHGMNALHVLEMANSPLNNGTEPGAFEGTVHRIEAKLTSVPKGKLP 300

Qy 241 TLLELHDYLNK1STVELEDFKYKELORLGUNNK1TDIENGSLANTIPRVRTHLENNKL 294

Db 235 TLLELHDENK1STVELEDKRYRELQIGLGNRITDENGTFANPRVRETHLENNKL 294

RESULT 8	
QY	301 KKIPSGLPELKYLQIIFLHSNISARYGVNDECPYVKMKSSAISLFFNPNPVKYEMQP 360
Db	295 KKPSGQELKYLQIIFLHSNISARYGVNDFCPTVKMKSSAISLFFNPNPVKYWEIOP 354
Qy	361 ATFRCVLSRMSVQLGNFG 378
Db	355 ATFRCVLGRMSVQLGNVG 372
RESULT 9	
ID	AAM40351 standard; Protein: 344 AA.
XX	
AC	AAM40351;
XX	
AC	
XX	
DT	22-OCT-2001 (first entry)
XX	
Human polypeptide SEQ ID NO 3496.	
Human: nootropic; immunosuppressant; cytostatic; gene therapy; cancer; peripheral nervous system; neuropathy; central nervous system; CNS; Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic; amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic; chemokinetic; thrombolytic; drug screening; arthritis; inflammation; leukaemia.	
XX	
OS	Homo sapiens.
XX	
PN	WO200153312-A1.
XX	
PD	26-JUL-2001.
XX	
PP	26-DEC-2000; 2000WO-US34263.
XX	
PR	21-JAN-2000; 2000US-0488725.
PR	25-APR-2000; 2000US-0552317.
PR	07-AUG-2000; 2000US-0598042.
PR	09-JUL-2000; 2000US-0598042.
PR	19-JUL-2000; 2000US-0620312.
PR	03-AUG-2000; 2000US-0653450.
PR	14-SEP-2000; 2000US-0662191.
PR	19-OCT-2000; 2000US-0693036.
PR	29-NOV-2000; 2000US-0727344.
XX	
PA	(HYSEQ INC.) HYSEQ INC.
XX	
Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D; Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA, Zhou P, Goodrich R, Drmanac RT;	
XX	
DR	WPI: 2001-442253/47.
XX	
PT	N-PSDB; AAI59507.
PT	
XX	Novel nucleic acids and polypeptides, useful for treating disorders such as central nervous system injuries -
PS	Example 6: SEQ ID NO 3496: 10070pp; English.
XX	
CC	The invention relates to human nucleic acids (AAI57798-AAI61369) and the encoded polypeptides (AAI38612-AAI42213) with nootropic, CC immunosuppressant and cytostatic activity. The polynucleotides are used CC in gene therapy. A composition containing a polypeptide or polynucleotide CC of the invention may be used to treat diseases of the peripheral nervous CC system, such as peripheral neuropathies and central nervous injuries, peripheral neuropathy and CC localised neuropathies and central nervous system diseases, such as CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the CC utilisation of the activities such as: Immune System suppression, CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic CC and thrombolytic activity, cancer diagnosis and therapy, drug screening CC assays for receptor activity, arthritis and inflammation, leukaemias and CC C.N.S disorders.
CC	Note: The sequence data for this patent did not form part of the printed CC specification.

XX	SQ	Sequence	344 AA:
	Query	Match	82.8%; Score 1649; DB 22; Length 344;
	Best Local Similarity	97.5%; Pred. No. 3..3e-145;	
	Matches 316; Conservative	1; Mismatches 7; Indels 0; Gaps 0	
Qy	56	FPTREPRSHFFPFEDLPMCPFGCQCYSSRVRHCSIDLGLTSVPTNIPEDTRMLDLQNNKIKE 115	
Db	21	FQEQSEAAFLFFDLPMPCPFGCQCYSSRVRHCSIDLGLTSVPTNIPEDTRMLDLQNNKIKE 80	
Qy	116	IKEENDFGLTSLYGLINNKTKIHPKAFLTTPKKLRLYLSHNOUSEIPNLPKSLAEL 175	
Db	81	IKEENDFGLTSLYGLINNKTKIHPKAFLTTPKKLRLYLSHNOUSEIPNLPKSLAEL 140	
Qy	176	RHENKVVKVTKIOTDTFKGMNALHLVLEMSANPLDNGTEPGAFEGTYVPHIRIAEAKLTSPV 235	
Db	141	RHENKVVKVTKIOTDTFKGMNALHLVLEMSANPLDNGTEPGAFEGTYVPHIRIAEAKLTSPV 200	
Qy	236	KGLPPTTLLELHDYNKNTISTVLEDFKRYKELQRQLGIGNKTDIENGSLANIPRVREIHL 295	
Db	201	KGLPPTTLLELHDYNKNTISTVLEDFKRYKELQRQLGIGNKTDIENGSLANIPRVREIHL 260	
Qy	296	ENNLKKKPSGLPELKYLQIIFLHSNSIARYGVNDCEPVPKMKSSLYSAISLFNPNPVKY 355	
Db	261	ENNLKKKPSGLPELKYLQIIFLHSNSIARYGVNDCEPVPKMKSSLYSAISLFNPNPVKY 320	
Qy	356	WEMQPATFRCVYLSRSMSVQLGNFGM 344	
Db	321	WEMQPATFRCVYLSRSMSVQLGNFGM 344	
RESULT 9			
	ABG22569		
ID	ABG22569	standard; protein; 352 AA.	
XX	AC	ABG22569;	
XX	DT	18-FEB-2002 (first entry)	
XX	DE	Novel human diagnostic protein #22560.	
XX	KW	Human; chromosome mapping; gene mapping; gene therapy; forensic	
XX	KW	food supplement; medical imaging; diagnostic; genetic disorder.	
OS	Homo sapiens.		
XX	PN	W0200175067-A2.	
XX	PD	11-OCT-2001.	
XX	PF	30-MAR-2001; 2001WO-US08631.	
XX	PR	31-MAR-2000; 2000US-0540217.	
XX	PR	23-AUG-2000; 2000US-0649167.	
PA	(HYSE-) HYSEQ INC.		
XX	PI	Drmancac RT, Liu C, Tang YT;	
XX	DR	WPI 2001-63962/73.	
XX	DR	N-PSDB; AAS86756.	
PT	New isolated polynucleotide and encoded polypeptides, useful in		
PT	diagnostics, forensics, gene mapping, identification of mutations		
PT	responsible for genetic disorders or other traits and to assess		
PT	biodiversity.		
XX	PS	Claim 20; SEQ ID No. 52928; 103pp; English.	
XX	CC	The invention relates to isolated polynucleotide (I) and	
CC	polypeptide (II) sequences. (I) is useful as hybridisation probes		

Page 8

and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (III). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. ABG0010-ABG3077 represent novel human diagnostic amino acid sequences of the invention.

Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at <http://www.wipo.int/patdb/notices.html>.

		Sequence	246 AA;
		Query Match	62.18;
		Best Local Similarity	96.78;
		Mismatches	0.5e-10;
		Matches	237;
		Conservative	1;
		Indels	3;
		Gaps	4;
		1 MKEYYLLFLALCSAKPFFPSPSHITALKNNMLKQMDT --- -DDDDDDDDDDDDDDNSLFLF 56	
		1 MKEYYLLFLALCSAKPFFPSPSHITALKNNMLKQMDT --- -DDDDDDDDDDDDDDNSLFLF 60	
Dy	57	PTRPRESHFPFPFDLPEMPCPFGCOCYSRVVHCSDGLITSVPTNIPFDTRMLDLQNNKTEI 116	
Db	61	PTRPRESHFPFPFDLPEMPCPFGCOCYSRVVHCSDGLITSVPTNIPFDTRMLDLQNNKTEI 120	
Dy	117	KENDFKGITSLYGLILNNNKLTKTHPKAFLTTKKLRRLYSHINQLSEPLNLPKSLAELR 176	
Db	121	KENDFKGITSLYGLILNNNKLTKTHPKAFLTTKKLRRLYSHINQLSEPLNLPKSLAELR 180	
Dy	177	THENKVKKIQKDTPKGMMALHVLEMSANPLDNGIEPGAFEGTVFHIRIAFAKLTSVPK 236	
Db	181	THENKVKKIQKDTPKGMMALHVLEMSANPLDNGIEPGAFEGTVFHIRIAFAKLTSVPK 240	
Dy	237	GLPPT 241	
		;	
Db	241	DNLPS 245	
	RESULT 11		
	AAR87951		
	AAU87951		
	AAR87951; standard; protein; 369 AA.		
	XX		
	AC		
	AAC		
	AAR87951;		

DT	20-MAR-1996	(first entry)	QY	300 LKKIPSGLPELKYLQIILHSNSNARYGVNDFCPTPKMKKSLSAISLFLNNPQVWEMO 359
XX	Rat neurotrophic biglycan.		Db	290 LSRFPAEGLDLKLQVVLHSNNNTKVGYINDCPMGRGKRAYNGTSLFNNPVPYWEVQ 349
KW	Biglycan; proteoglycan; chondroitin sulphate; neuron protection;		QY	360 PATRCVCILSRMSVOLGNF 377
KW	neurotrophic; central nervous system; CNS; memory loss; dementia;		Db	350 PATRCVCILRBLAQFGNY 367
XX	learning.			
OS	Rattus sp.			
XX				
FH				
Key			RESULT 12	
Peptide			ID AAB85043	
FT			AA85043 standard; Protein; 368 AA.	
Region			XX	
FT			AC AAB85043;	
XX			XX	
W09530432-A1.			DT 06-AUG-2001 (first entry)	
1..37			XX	
/label= sig_peptide			DE Human biglycan protein sequence.	
44..60			XX	
/label= Hypervariable_region			Dystrophin-associated protein complex; DAPC; postsynaptic membrane;	
XX			KW biglycan; muscular dystrophy; neuromuscular; neurological; smooth muscle;	
XX			KW nootropic; neuroleptic; antiviral.	
XX			XX	
XX			OS Homo sapiens.	
XX			XX	
XX			PN WO200136475-A2.	
XX			XX	
DR			PD 25-MAY-2001.	
N-PSDB; AAT0876.			XX	
XX			PF 17 NOV-2000; 2000WO-US31661.	
XX			XX	
WPI: 1995-403938/51.			PR 18-NOV-1999; 99US-0166253.	
DR			XX	
XX			(YBR-) UNTV BROWN RES FOUND.	
XX			XX	
09-MAY-1994;		94WO-EP01479.	PI Fallon J, McKechnie B, Rafi M, Creely H, Bowe M, Ferri R;	
XX			XX	
PR			DR WPI: 2001-355617/37.	
XX			XX	
(BOEF ) BOEHRINGER MANNHEIM GMBH.			DR N-PSDB; AAT83977.	
XX			XX	
PI Hasenoehrl R, Huston J, Junghans U, Kappeler J, Koops A;			Stabilizing dystrophin-associated protein complexes and activating	
PI Mueller HW;			PT postsynaptic membrane of a cell for treating or preventing muscular,	
XX			PT neuromuscular and neurological disorders, involves contacting cell with	
WPI: 1995-403938/51.			PT biglycan.	
DR			XX	
N-PSDB; AAT0876.			PS Disclosure; Page 101-108; 112pp; English.	
XX			XX	
XX			CC The invention relates to stabilizing dystrophin-associated protein	
XX			CC complexes (DAPCs) on the surface of a cell or activating a postsynaptic	
XX			CC membrane of a cell that comprises contacting the cell with an effective	
XX			CC amount of biglycan. A composition comprising biglycan or its portion	
XX			CC is useful for treating or preventing a condition associated with abnormal	
XX			CC DAPC in cells, characterized by breakdown of muscle cell membrane, which	
XX			CC includes muscular dystrophies, such as Duchenne's, Becker's, Congenital,	
XX			CC Limb-girdle muscular dystrophy and myotonic dystrophy and a condition	
XX			CC characterized by abnormal neuromuscular junction or synapse, such as	
XX			CC neuromuscular or neurological diseases in a subject. Neurological	
XX			CC disorders, include polymyositis and Alzheimer's disease. Biglycan is also	
XX			CC useful for preventing and treating smooth muscle disorders, such as	
XX			CC cardiac myopathies and for treating and inhibiting infections of cells by	
XX			CC microorganisms e.g. viruses. Agents that modulate the activity of	
XX			CC DAG-125, identified by the methods are useful in the prophylactic and	
XX			CC therapeutic treatments of diseases or disorders, characterized by an	
XX			CC unstable DAPC or an inappropriate formation of a postsynaptic	
XX			CC differentiation. Biglycans are also useful as a supplement to brain or	
XX			CC muscle cell or tissue culture and, tissues can be incubated in vitro	
XX			CC with biglycan to reverse tissue atrophy and to improve their growth or	
XX			CC survival in vitro. The present sequence represents the human biglycan.	
XX			XX	
Sequence	369 AA;		SQ Sequence 368 AA;	
Query Match	52.2%	Score 1040.5; DB 16; Length 369;	Query Match	51.98; Score 1034; DB 22; Length 368;
Best Local Similarity	52.6%	Pred. No. 2e-88;	Best Local Similarity	52.6%
Matches	199; Conservative	Mismatches 93; Indels 19; Gaps 3;	Matches	199; Conservative
QY	4 YVLLFLALCSAKPE ---FSSPHIAKNNMLKDMEIDDDDDDDDDDDNSLFPTR 59		QY	4 YVLLFLALCSAKPE ---FSSPHIAKNNMLKDMEIDDDDDDDDDDDNSLFPTR 59
Db	5 WLTULLALSAOLPPEQKGFWDFLDDGLIMMDEEASGSDTSGVLDL---SLPT-	59	Db	5 WLTULLALSAOLPPEQKGFWDFLDDGLIMMDEEASGSDTSGVLDL---SLPT-
QY	60 EPRSHFPFDIFPMCPFGCGCYSSRVHCSDLGLTSVSPNIPDTRMLDNQNIKEIKEN 119		QY	60 EPRSHFPFDIFPMCPFGCGCYSSRVHCSDLGLTSVSPNIPDTRMLDNQNIKEIKEN 119
Db	60 -----FSAMCPFGCHLURVQCSDLGKTVPREISDPTLUDLQNDNSEIRKD 109		Db	60 -----FSAMCPFGCHLURVQCSDLGKTVPREISDPTLUDLQNDNSEIRKD 109
QY	120 DEFGLTSLYGLILNNNNKLTKTHPKAELTTKLRLLYLSHNQSEIPNLPKSLAELRIHE 179		QY	120 DEFGLTSLYGLILNNNNKLTKTHPKAELTTKLRLLYLSHNQSEIPNLPKSLAELRIHE 179
Db	110 DFKGLOHLYAVLNKNISKIHEKAFLPLQLKLYLVEIPNLSSLVEIRHD 169		Db	110 DFKGLOHLYAVLNKNISKIHEKAFLPLQLKLYLVEIPNLSSLVEIRHD 169
QY	180 NVKKVIQDKTFKGMLAHYLEMSANPLDNGTEPGAFRGTVFHIRIAEKLTSPKGLP 239		QY	180 NVKKVIQDKTFKGMLAHYLEMSANPLDNGTEPGAFRGTVFHIRIAEKLTSPKGLP 239
Db	170 NRKVPKGFSGLRNMCNTMGENPLENSGPAGDGLKLNYLRSEAKLTGPKDLP 229		Db	170 NRKVPKGFSGLRNMCNTMGENPLENSGPAGDGLKLNYLRSEAKLTGPKDLP 229
Qy	240 PTLELHLDDNKTISTKQIATELDLRYSKLYRLGQHQNTRMENGSLSLFPLTLREHLDNNK 299		Qy	240 PTLELHLDDNKTISTKQIATELDLRYSKLYRLGQHQNTRMENGSLSLFPLTLREHLDNNK 299
Db	230 ETLNELHLDDNKTQATELDLRYSKLYRLGQHQNTRMENGSLSLFPLTLREHLDNNK 289		Db	230 ETLNELHLDDNKTQATELDLRYSKLYRLGQHQNTRMENGSLSLFPLTLREHLDNNK 289



PR 08-SEP-1994; 94US-0303238.  
 PR 28-JUN-1988; 88US-0212702.  
 PR 22-JAN-1990; 90US-0467884.  
 PR 13-MAY-1992; 92US-0882345.  
 PA (BURN-) BURNHAM INST.  
 XX  
 PI Ruoslahti EI, Yamaguchi Y;  
 XX WPI: 2001-610491/70.  
 XX  
 PT Use of decorin or biglycan polypeptide for inhibiting TGF-beta activity  
 PT in the treatment of dermal wounds and cancer -  
 XX  
 PS Example 7: Fig 11: 40pp; English.  
 XX  
 CC The invention relates to the inhibition of transforming growth factor-beta (TGF-beta) activity involving contacting TGF-beta with a purified polypeptide comprising leucine-rich amino acid sequence of a member of decorin superfamily of mammalian proteoglycans. The following activities can be attributed to the polypeptide of the invention:  
 CC cytostatic, nephrotropic, antirheumatic, antiarthritic, vasotropic,  
 CC antiarteriosclerotic, hepatotropic, cardiotropic, dermatological and  
 CC pulmonary. Polypeptides of the invention act as transforming growth  
 CC factor-beta (TGF-beta) binder. The polypeptides of the invention can be used for treating a pathology, particularly proliferative pathology caused by a transforming growth factor-beta (TGF-beta) regulated activity such as cancer; particularly fibrotic cancer, fibrotic disease, glomerulonephritis, rheumatoid arthritis, arteriosclerosis, adult respiratory distress syndrome, cirrhosis of liver, fibrosis of lungs, post-myocardial infarction, cardiac fibrosis, post-angioplasty restenosis, renal interstitial fibrosis and certain dermal fibrotic conditions such as keloids and scarring resulting from burn injuries; other invasive skin injuries and reconstructive surgery. The wounds treated with the polypeptide, particularly decorin exhibit no detectable scarring, and are histologically normal. The current sequence represents human biglycan.  
 XX Sequence 368 AA;

Query Match 51.3%; Score 1021; DB 22; Length 368;  
 Best Local Similarity 52.4%; Pred. No. 1.3e-86;  
 Matches 197; Conservative 64; Mismatches 95; Indels 20; Gaps 3;

Qy . 6 LLLFLALCSAKPF---FSPSHITALKNNMKDMEIDDDDDDDDDDDDNNSLEPTREP 61  
 . 7 LVSLAISQALFEQRGFWDFLTDQGPMMNDEEASGADTSVQLPDPD----SVPTYS- 60

Qy 62 RSHFFPDLFPMCPFGCQCYSRVHCSDIGLTSVTPNTIPDTRMLDQNKKIKEIKENDF 121  
 Db 61 -----AMCPFGCHCIRVYQCSLIGKSVKEISDPTTLQDNNDISEURKDDF 110

Qy 122 KGLTSLYGLINNNKLTKHPKAFLDTKLRRLYLSHNOSEIPLNLPKSLAELRHENK 181  
 Db 111 KLGQHLYALVNNKLISKHEKAFLSPRNQKLYLSKNHLVEPPNPLSSVLERIHDR 170

Qy 182 VRIKIQDKTFKGMDNLHVLEMSANPLDNGTEPAGFEGTVFHIRTAEAKTSWPKGLPPT 241  
 Db 171 IRKVPGKVSGLRNMNCIEMGGNPDLNSGEPGADGLKLYLRSEAKTGPKDLPET 230

Qy 242 LLELHDYDKLSTVLEDFKRYKELORLIGGNKLTDIENGSLANIPLYREIHNENKLK 301  
 Db 231 LNELHDHNIKQIAIEEDLLRYSKLYRLGLGHQTRMENGSLSFPLTRELHDNNKLA 290

Qy 302 KIPSGLPELYKQIIFLHSNSIARYGVNDEFCTPVPKMKSLYSAISLFLNPVAKWEMOPA 361  
 Db 291 RVPGLGLPDLKLUQVYVLSNSNTKVGVDNFCPMEGFGVKRAYINGTSLEFNPNPVWEVODA 350

Qy 362 TFCRVLSRMSVQLGNE 377  
 . 351 TFRCVTDRLLAIQFGNY 366

Qy 422 LLELHDYDKLSTVLEDFKRYKELORLIGGNKLTDIENGSLANIPLYREIHNENKLK 301  
 Db 231 LNELHDHNIKQIAIEEDLLRYSKLYRLGLGHQTRMENGSLSFPLTRELHDNNKLA 290

RESULT\_15  
 AAR05159 standard; protein: 368 AA.  
 ID AAR05159;  
 XX  
 AC AAR05159;  
 XX DT 17-DEC-2001 (updated)  
 DT 09-OCT-1990 (first entry)  
 XX DE Sequence of human bone proteoglycan I (biglycan).  
 XX KW Osteoporosis; rheumatoid arthritis; Pager's disease;  
 KW atherosclerosis; periodontal; human bone matrix; proteoglycan.  
 XX OS Homo sapiens.  
 XX PN USN7432044\_N.  
 XX PD 17-APR-1990.  
 XX PP 03-NOV-1989;  
 XX PR 03-NOV-1989;  
 XX PA (US SH ) NAT INST OF HEALTH.  
 PI Termino J;  
 XX WPI: 1990-178641/23.  
 DR N-PSDB; AAQ04490.

CC Human bone matrix DNA and proteins -  
 CC PT Human in detection, diagnosis and treatment involving skeletal  
 PT and/or connective tissue disease states.  
 XX Disclosure: ; p: English.  
 XX  
 CC Probes and Abs raised to the proteins can be used to determine  
 CC their levels useful in diagnosis of associated connective tissue  
 CC diseases states such as osteoporosis, osteorheumatoid arthritis,  
 CC Proteins may also be used to induce or block biological function.  
 CC (Note: Revised entry submitted to correct the patent number format of  
 CC US Government-owned NTIS applications to prevent clashes with ongoing US  
 CC web site at www.dowdow.com/dwp1/updates/ntis\_us.htm.)  
 XX SQ Sequence 368 AA;

Query Match 51.2%; Score 1020; DB 11; Length 368;  
 Best Local Similarity 52.1%; Pred. No. 1.6e-86;  
 Matches 196; Conservative 66; Mismatches 94; Indels 20; Gaps 3;

Qy 6 LLLFLALCSAKPF---FSPSHITALKNNMKDMEIDDDDDDDDDDNNSLEPTREP 61  
 . 7 LVSLAISQALFEQRGFWDFLTDQGPMMNDEEASGADTSVQLPDPD----SVPTYS- 60

Qy 122 KGLTSLYGLINNNKLTKHPKAFLDTKLRRLYLSHNOSEIPLNLPKSLAELRHENK 181  
 Db 61 -----AMCPFGCHCIRVYQCSLIGKSVKEISDPTTLQDNNDISEURKDDF 110

Qy 182 VRIKIQDKTFKGMDNLHVLEMSANPLDNGTEPAGFEGTVFHIRTAEAKTSWPKGLPPT 241  
 Db 171 IRKVPGKVSGLRNMNCIEMGGNPDLNSGEPGADGLKLYLRSEAKTGPKDLPET 230

Qy 242 LLELHDYDKLSTVLEDFKRYKELORLIGGNKLTDIENGSLANIPLYREIHNENKLK 301  
 Db 231 LNELHDHNIKQIAIEEDLLRYSKLYRLGLGHQTRMENGSLSFPLTRELHDNNKLA 290

Qy 302 KIPSGLPELYKQIIFLHSNSIARYGVNDEFCTPVPKMKSLYSAISLFLNPVAKWEMOPA 361  
 Db 291 RVPGLGLPDLKLUQVYVLSNSNTKVGVDNFCPMEGFGVKRAYINGTSLEFNPNPVWEVODA 350

Qy 362 TFCRVLSRMSVQLGNE 377  
 . 351 TFRCVTDRLLAIQFGNY 366

Qy 422 LLELHDYDKLSTVLEDFKRYKELORLIGGNKLTDIENGSLANIPLYREIHNENKLK 301  
 Db 231 LNELHDHNIKQIAIEEDLLRYSKLYRLGLGHQTRMENGSLSFPLTRELHDNNKLA 290

291	RVPGLPDLKLQVYLLHSNNITKVGVNDFCPMGFGVRAYINGTSLFNNPVPWYEQPA	350
292	KIPGLPPELKYLQIIFLHSNSTARVGVNDFCPTWPKMKSLISATISLFNNPVPWYEQPA	361
293	:     :    :    :    :    :    :    :    :    :    :    :    :	
362	TFRCVLRSMSVQLGNF	377
351	TFRCVTDRLAFOGNY	366

Search completed: May 27, 2003, 16:04:07  
Job time : 39 secs

GenCore version 5.1.4\_P5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: May 27, 2003, 16:04:11 ; Search time 23 Seconds  
(without alignments)  
1634.193 Million cell updates/sec

Title: US-09-944-457-2  
Perfect score: 1932

Sequence: 1 NKEVLLFLALCSAKPFFS.....PATFRCYVLRSRMSVQLGNFGM 379

Scoring table: BLOSUM62  
Gap0 10.0 , Gapext 0.5

Archived: 375593 seqs, 99172665 residues

Total number of hits satisfying chosen parameters: 375593

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:\*

1: /cgns2\_6/podata/2/pubpa/US08\_NEWPUB.pep;\*  
2: /cgns2\_6/podata/2/pubpa/PCT\_NEWPUB.pep;\*  
3: /cgns2\_6/podata/2/pubpa/US06\_NEWPUB.pep;\*  
4: /cgns2\_6/podata/2/pubpa/US05\_NEWPUB.pep;\*  
5: /cgns2\_6/podata/2/pubpa/US07\_NEWPUB.pep;\*  
6: /cgns2\_6/podata/2/pubpa/US01\_PUBCOMB.pep;\*  
7: /cgns2\_6/podata/2/pubpa/PCTUS\_PUBCOMB.pep;\*  
8: /cgns2\_6/podata/2/pubpa/US08\_PUBCOMB.pep;\*  
9: /cgns2\_6/podata/2/pubpa/US09\_NEWPUB.pep;\*  
10: /cgns2\_6/podata/2/pubpa/US09\_PUBCOMB.pep;\*  
11: /cgns2\_6/podata/2/pubpa/US10\_NEWPUB.pep;\*  
12: /cgns2\_6/podata/2/pubpa/US10\_PUBCOMB.pep;\*  
13: /cgns2\_6/podata/2/pubpa/US60\_NEWPUB.pep;\*  
14: /cgns2\_6/podata/2/pubpa/US60\_PUBCOMB.pep;\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

## 8

Result No.	Score	Match Length	DB ID	Description
1	1992	100.0	379	9 US-09-944-413-2
2	1992	100.0	379	9 US-09-944-403-2
3	1992	100.0	379	9 US-09-944-896-2
4	1992	100.0	379	9 US-09-944-444-2
5	1992	100.0	379	9 US-09-944-907-2
6	1992	100.0	379	9 US-09-944-929-2
7	1992	100.0	379	9 US-10-028-072-128
8	1992	100.0	379	9 US-10-121-049-128
9	1992	100.0	379	9 US-10-123-904-128
10	1992	100.0	379	9 US-10-140-070-128
11	1992	100.0	379	9 US-10-175-746-328
12	1992	100.0	379	9 US-10-176-918-328
13	1992	100.0	379	9 US-10-176-121-128
14	1992	100.0	379	9 US-10-137-865-128
15	1992	100.0	379	9 US-10-140-474-328
16	1992	100.0	379	9 US-10-142-431-328
17	1992	100.0	379	9 US-10-143-114-328
18	1992	100.0	379	9 US-10-140-002-328
19	1992	100.0	379	9 US-10-142-419-328

Result No.	Score	Match Length	DB ID	Description
1	1992	100.0	379	9 US-10-123-262-328
2	1992	100.0	379	9 US-10-142-423-328
22	1992	100.0	379	9 US-10-121-050-328
23	1992	100.0	379	9 US-10-141-755-328
24	1992	100.0	379	9 US-10-143-032-328
25	1992	100.0	379	9 US-10-123-108-328
26	1992	100.0	379	9 US-10-123-236-328
27	1992	100.0	379	9 US-10-123-261-328
28	1992	100.0	379	9 US-10-140-921-328
29	1992	100.0	379	9 US-10-140-926-328
30	1992	100.0	379	9 US-10-121-045-328
31	1992	100.0	379	9 US-10-123-292-328
32	1992	100.0	379	9 US-10-123-903-328
33	1992	100.0	379	9 US-10-121-041-328
34	1992	100.0	379	9 US-10-121-047-328
35	1992	100.0	379	9 US-10-140-925-328
36	1992	100.0	379	9 US-10-160-498-328
37	1992	100.0	379	9 US-09-944-884-2
38	1992	100.0	379	9 US-10-121-041-328
39	1992	100.0	379	9 US-10-121-047-328
40	1992	100.0	379	9 US-10-121-047-328
41	1992	100.0	379	9 US-10-123-215-328
42	1992	100.0	379	9 US-10-123-902-328
43	1992	100.0	379	9 US-10-123-908-328
44	1992	100.0	379	9 US-10-123-909-328
45	1992	100.0	379	9 US-10-123-910-328

## ALIGNMENTS

RESULT 1 US-09-944-413-2					
Sequence 2	Application US/09944413				
Patent No.	US20020156004A1				
GENERAL INFORMATION:					
APPLICANT:	Baker, Kevin				
APPLICANT:	Botstein, David				
APPLICANT:	Eaton, Dan				
APPLICANT:	Ferrara, Napoleone				
APPLICANT:	Filvaroff, Ellen				
APPLICANT:	Gerritsen, Mary				
APPLICANT:	Goddard, Audrey				
APPLICANT:	Godowski, Paul				
APPLICANT:	Grimaldi, Christopher				
APPLICANT:	Gurney, Austin				
APPLICANT:	Hillman, Kenneth				
APPLICANT:	Klijavin, Ivar				
APPLICANT:	Napier, Mary				
APPLICANT:	Roy, Margaret				
APPLICANT:	Tumas, Daniel				
APPLICANT:	Wood, William				
TITLE OF INVENTION:	SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC ACIDS ENCODING THE SAME				
FILE REFERENCE:	P2548P1C1				
CURRENT APPLICATION NUMBER:	US/09/944,413				
CURRENT FILING DATE:	2001-09-26				
PRIOR APPLICATION NUMBER:	09/816,028				
PRIOR FILING DATE:	2001-05-25				
PRIOR APPLICATION NUMBER:	60/067,411				
PRIOR FILING DATE:	December 3, 1997				
PRIOR APPLICATION NUMBER:	60/069,334				
PRIOR FILING DATE:	December 11, 1997				
PRIOR APPLICATION NUMBER:	60/069,335				
PRIOR FILING DATE:	December 11, 1997				
PRIOR APPLICATION NUMBER:	60/069,278				
PRIOR FILING DATE:	December 11, 1997				
PRIOR APPLICATION NUMBER:	60/069,425				
PRIOR FILING DATE:	December 12, 1997				
PRIOR APPLICATION NUMBER:	60/069,616				
PRIOR FILING DATE:	December 16, 1997				
PRIOR APPLICATION NUMBER:	60/069,694				
PRIOR FILING DATE:	December 16, 1997				

PRIOR APPLICATION NUMBER: 60/069,702  
 PRIOR FILING DATE: December 16, 1997  
 PRIOR APPLICATION NUMBER: 60/069,870  
 PRIOR FILING DATE: December 17, 1997  
 PRIOR APPLICATION NUMBER: 60/069,873  
 PRIOR FILING DATE: December 17, 1997  
 PRIOR APPLICATION NUMBER: 60/068,017  
 PRIOR FILING DATE: December 18, 1997  
 PRIOR APPLICATION NUMBER: 60/070,440  
 PRIOR FILING DATE: January 5, 1998  
 PRIOR APPLICATION NUMBER: 60/074,086  
 PRIOR FILING DATE: February 9, 1998  
 PRIOR APPLICATION NUMBER: 60/074,092  
 PRIOR FILING DATE: February 9, 1998  
 PRIOR APPLICATION NUMBER: 60/075,945  
 PRIOR FILING DATE: February 25, 1998  
 PRIOR APPLICATION NUMBER: 60/112,850  
 PRIOR FILING DATE: December 16, 1998  
 PRIOR APPLICATION NUMBER: 60/113,296  
 PRIOR FILING DATE: December 22, 1998  
 PRIOR APPLICATION NUMBER: 60/146,222  
 PRIOR FILING DATE: July 28, 1999  
 PRIOR APPLICATION NUMBER: PCT/US98/19330  
 PRIOR FILING DATE: September 16, 1998  
 PRIOR APPLICATION NUMBER: PCT/US98/25108  
 PRIOR FILING DATE: December 1, 1998  
 PRIOR APPLICATION NUMBER: 09/216,021  
 PRIOR FILING DATE: December 16, 1998  
 PRIOR APPLICATION NUMBER: 09/218,517  
 PRIOR FILING DATE: December 22, 1998  
 PRIOR APPLICATION NUMBER: 09/254,311  
 PRIOR FILING DATE: March 3, 1999  
 PRIOR APPLICATION NUMBER: PCT/US99/12252  
 PRIOR FILING DATE: June 22, 1999  
 PRIOR APPLICATION NUMBER: PCT/US99/21090  
 PRIOR FILING DATE: September 15, 1999  
 PRIOR APPLICATION NUMBER: PCT/US99/28409  
 PRIOR FILING DATE: NO. US2002156004A1ember 30, 1999  
 PRIOR APPLICATION NUMBER: PCT/US99/28313  
 PRIOR FILING DATE: NO. US2002156004A1ember 30, 1999  
 PRIOR APPLICATION NUMBER: PCT/US99/30095  
 PRIOR FILING DATE: December 16, 1999  
 PRIOR APPLICATION NUMBER: PCT/US00/03565  
 PRIOR FILING DATE: February 11, 2000  
 PRIOR APPLICATION NUMBER: PCT/US00/04414  
 PRIOR FILING DATE: February 22, 2000  
 PRIOR APPLICATION NUMBER: PCT/US00/05841  
 PRIOR FILING DATE: March 2, 2000  
 PRIOR APPLICATION NUMBER: PCT/US00/08439  
 PRIOR FILING DATE: March 30, 2000  
 PRIOR FILING DATE: May 22, 2000  
 PRIOR APPLICATION NUMBER: PCT/US00/20110  
 PRIOR FILING DATE: July 28, 2000  
 PRIOR APPLICATION NUMBER: PCT/US00/32678  
 PRIOR FILING DATE: December 1, 2000  
 PRIOR APPLICATION NUMBER: PCT/US01/06520  
 NUMBER OF SEQ ID NOS: 120  
 SEQ ID NO 2  
 LENGTH: 379  
 TYPE: PRT  
 ORGANISM: Homo Sapien  
 US-09-944-413-2

Query Match Score 1992; DB 9; Length 379;  
 Best Local Similarity 100.0%; Pred. No. 8. 6e-159;  
 Matches 379; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MKEYVLLFLALCSAKPFFSPSHIALKNNMLKMDTDDDDDEDSNLSLEPTRE 60  
 1 MKEYVLLFLALCSAKPFFSPSHIALKNNMLKMDTDDDDDEDSNLSLEPTRE 60  
 1 PRSHFEPFDLPEPMCPEGCQCYSRVHCSDIGLTSYPTNLPEFDTRMLDNKIKETKEND 120  
 1 PRSHFEPFDLPEPMCPEGCQCYSRVHCSDIGLTSYPTNLPEFDTRMLDNKIKETKEND 120  
 1 FKGLTSLYGLLNNNNKLTKTHPKAFLTTKLRLYLSSHQNOLSEIPUNLPKSLAELRITHEN 180  
 1 FKGLTSLYGLLNNNNKLTKTHPKAFLTTKLRLYLSSHQNOLSEIPUNLPKSLAELRITHEN 180  
 1 FKGLTSYGLLNNNNKLTKTHPKAFLTTKLRLYLSSHQNOLSEIPUNLPKSLAELRITHEN 180  
 1 FKGLTSYGLLNNNNKLTKTHPKAFLTTKLRLYLSSHQNOLSEIPUNLPKSLAELRITHEN 180  
 1 KVKKIQKDTEFGMNAJHVLENSANPLDNNNGEPAFEGTVFHIRAEAKLTSPKGGLPP 240  
 1 KVKKIQKDTEFGMNAJHVLENSANPLDNNNGEPAFEGTVFHIRAEAKLTSPKGGLPP 240  
 1 KVKKIQKDTEFGMNAJHVLENSANPLDNNNGEPAFEGTVFHIRAEAKLTSPKGGLPP 240  
 1 TLEELHDYWKNTSTVELEDKRYKELQRGIGNNKTIDINGSLANTPRYREHLENKL 300  
 1 TLEELHDYWKNTSTVELEDKRYKELQRGIGNNKTIDINGSLANTPRYREHLENKL 300  
 1 KKPSGLPELKYLQLIPLFHNSIARYGVNDFCPTVPKMKSLYSALSLENNPVKYWEMQP 360  
 1 KKPSGLPELKYLQLIPLFHNSIARYGVNDFCPTVPKMKSLYSALSLENNPVKYWEMQP 360  
 1 ATFRCVLSRMSVQLGNFGM 379  
 1 ATFRCVLSRMSVQLGNFGM 379  
 1 ATFRCVLSRMSVQLGNFGM 379  
 1 ATFRCVLSRMSVQLGNFGM 379  
 RESULT 2  
 US-09-944-403-2  
 ; Sequence 2, Application US/09944403  
 ; Patent No. US20020165143A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Baker, Kevin  
 ; APPLICANT: Boistein, David  
 ; APPLICANT: Eaton, Dan  
 ; APPLICANT: Ferrara, Napoleone  
 ; APPLICANT: Filvaroff, Ellen  
 ; APPLICANT: Gerritsen, Mary  
 ; APPLICANT: Goddard, Audrey  
 ; APPLICANT: Goodwski, Paul  
 ; APPLICANT: Grimaldi, Christopher  
 ; APPLICANT: Gurley, Austin  
 ; APPLICANT: Hillin, Kenneth  
 ; APPLICANT: Kjavian, Ivar  
 ; APPLICANT: Napier, Mary  
 ; APPLICANT: Roy, Margaret  
 ; APPLICANT: Tunas, Daniel  
 ; APPLICANT: Wood, William  
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
 ; ACIDS ENCODING THE SAME  
 ; FILE REFERENCE: P2544P1C1  
 ; CURRENT APPLICATION NUMBER: US/09/944,403  
 ; CURRENT FILING DATE: 2001-09-26  
 ; PRIOR APPLICATION NUMBER: 09/866,028  
 ; PRIOR FILING DATE: 2001-05-25  
 ; PRIOR APPLICATION NUMBER: 60/067,411  
 ; PRIOR FILING DATE: December 3, 1997  
 ; PRIOR APPLICATION NUMBER: 60/069,334  
 ; PRIOR FILING DATE: December 11, 1997  
 ; PRIOR APPLICATION NUMBER: 60/069,335  
 ; PRIOR FILING DATE: December 11, 1997  
 ; PRIOR APPLICATION NUMBER: 60/069,278  
 ; PRIOR FILING DATE: December 11, 1997  
 ; PRIOR APPLICATION NUMBER: 60/069,425  
 ; PRIOR FILING DATE: December 12, 1997  
 ; PRIOR APPLICATION NUMBER: 60/069,696  
 ; PRIOR FILING DATE: December 16, 1997  
 ; PRIOR APPLICATION NUMBER: 60/069,694  
 ; PRIOR FILING DATE: December 16, 1997  
 ; PRIOR APPLICATION NUMBER: 60/069,702  
 ; PRIOR FILING DATE: December 16, 1997  
 ; PRIOR APPLICATION NUMBER: 60/069,870  
 ; PRIOR FILING DATE: December 17, 1997

Query Match Score 1992; DB 9; Length 379;  
 Best Local Similarity 100.0%; Pred. No. 8. 6e-159;  
 Matches 379; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MKEYVLLFLALCSAKPFFSPSHIALKNNMLKMDTDDDDDEDSNLSLEPTRE 60  
 1 MKEYVLLFLALCSAKPFFSPSHIALKNNMLKMDTDDDDDEDSNLSLEPTRE 60

PRIOR APPLICATION NUMBER: 60/069, 873  
; PRIOR FILING DATE: December 17, 1997  
; PRIOR APPLICATION NUMBER: 60/068, 017  
; PRIOR FILING DATE: December 18, 1997  
; PRIOR APPLICATION NUMBER: 60/070, 440  
; PRIOR FILING DATE: January 5, 1998  
; PRIOR APPLICATION NUMBER: 60/074, 086  
; PRIOR FILING DATE: February 9, 1998  
; PRIOR APPLICATION NUMBER: 60/074, 092  
; PRIOR FILING DATE: February 9, 1998  
; PRIOR APPLICATION NUMBER: 60/075, 945  
; PRIOR FILING DATE: February 25, 1998  
; PRIOR APPLICATION NUMBER: 60/112, 850  
; PRIOR FILING DATE: December 16, 1998  
; PRIOR APPLICATION NUMBER: 60/113, 236  
; PRIOR FILING DATE: December 22, 1998  
; PRIOR APPLICATION NUMBER: 60/146, 222  
; PRIOR FILING DATE: July 28, 1999  
; PRIOR APPLICATION NUMBER: PCT/US98/19330  
; PRIOR FILING DATE: September 16, 1998  
; PRIOR APPLICATION NUMBER: PCT/US98/25108  
; PRIOR FILING DATE: December 1, 1998  
; PRIOR APPLICATION NUMBER: 09/216, 021  
; PRIOR FILING DATE: December 16, 1998  
; PRIOR APPLICATION NUMBER: 09/218, 517  
; PRIOR FILING DATE: December 22, 1998  
; PRIOR APPLICATION NUMBER: 09/254, 311  
; PRIOR FILING DATE: March 3, 1999  
; PRIOR APPLICATION NUMBER: PCT/US99/112252  
; PRIOR FILING DATE: June 22, 1999  
; PRIOR APPLICATION NUMBER: PCT/US99/21190  
; PRIOR FILING DATE: September 15, 1999  
; PRIOR APPLICATION NUMBER: PCT/US99/284109  
; PRIOR FILING DATE: No. US20020165143A1ember 30, 1999  
; PRIOR APPLICATION NUMBER: PCT/US99/28413  
; PRIOR FILING DATE: No. US20020165143A1ember 30, 1999  
; PRIOR APPLICATION NUMBER: PCT/US99/28401  
; PRIOR FILING DATE: December 1, 1999  
; PRIOR APPLICATION NUMBER: PCT/US99/30095  
; PRIOR FILING DATE: December 16, 1999  
; PRIOR APPLICATION NUMBER: PCT/US00/03565  
; PRIOR FILING DATE: March 30, 2000  
; PRIOR APPLICATION NUMBER: PCT/US00/14042  
; PRIOR FILING DATE: May 22, 2000  
; PRIOR APPLICATION NUMBER: PCT/US00/05841  
; PRIOR FILING DATE: March 2, 2000  
; PRIOR APPLICATION NUMBER: PCT/US00/08439  
; PRIOR FILING DATE: March 30, 2000  
; PRIOR APPLICATION NUMBER: PCT/US00/14042  
; PRIOR FILING DATE: May 22, 2000  
; PRIOR FILING DATE: February 22, 2000  
; PRIOR FILING DATE: December 1, 1999  
; PRIOR FILING DATE: July 28, 2000  
; PRIOR APPLICATION NUMBER: PCT/US00/32678  
; PRIOR FILING DATE: December 1, 2000  
; PRIOR APPLICATION NUMBER: PCT/US01/06520  
; PRIOR FILING DATE: February 28, 2001  
; NUMBER OF SEQ ID NOS: 120  
; SEQ ID NO 2  
; LENGTH: 379  
; TYPE: PRT  
; ORGANISM: Homo Sapien

Query Match 100.0% Score 1992: DB 9; Length 379;  
Best Local Similarity 100.0% Pred. No. 8.6e-15%;  
Matches 379; Conservative 0; Mismatches 0; Gaps 0;

Db 61 PRSHFPFDLFPMPFGCQCYSRVHCSDLGLTSVPNIPFDTRMLDQNNNIKEIKEND 120  
Qy 1 MKEYVLLFLALCSAKPFSPSHALKNMKMDTDDDDDDDDDENSLEPTRE 60  
Db 1 MKEYVLLFLALCSAKPFSPSHALKNMKMDTDDDDDDDDDENSLEPTRE 60  
Qy 61 PRSHFPFDLFPMPFGCQCYSRVHCSDLGLTSVPNIPFDTRMLDQNNNIKEIKEND 120

Db 61 PRSHFPFDLFPMPFGCQCYSRVHCSDLGLTSVPNIPFDTRMLDQNNNIKEIKEND 120  
Qy 121 FKGLTSLYGLTINNNKKTKHPLAFLLTPIKLLRLLYLSHNOLESETPLNPKSLAELRTHEN 180  
Db 121 FKGLTSLYGLTINNNKKTKHPLAFLLTPIKLLRLLYLSHNOLESETPLNPKSLAELRTHEN 180  
Qy 181 KVKKIQDKTFKGMMALHYLEMMSANPLDONGTEPGAFEGTVFHITAEAKLTSVKGLPP 240  
Db 181 KVKKIQDKTFKGMMALHYLEMMSANPLDONGTEPGAFEGTVFHITAEAKLTSVKGLPP 240  
Qy 241 TLEFLHDYLNKISIVELDEFKYLQKELORLGLGNKFTDTEGSLANIPREIHNKL 300  
Db 241 TLEFLHDYLNKISIVELDEFKYLQKELORLGLGNKFTDTEGSLANIPREIHNKL 300  
Qy 301 KKPSGLDEKLQLIETHNSNARVEVDNDCPVTKMKSSLYSAISLFLNPVQYWEMQP 360  
Db 301 KKPSGLDEKLQLIETHNSNARVEVDNDCPVTKMKSSLYSAISLFLNPVQYWEMQP 360  
Qy 361 ATPERCVLSRMSVOLGNFGM 379  
Db 361 ATPERCVLSRMSVOLGNFGM 379

RESULT 3  
US-09-944-896-2  
; Sequence 2, Application US/09944896  
; Patent No. US20020168715A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin  
; APPLICANT: Botstein, David  
; APPLICANT: Eatan, Dan  
; APPLICANT: Ferrara, Napoleone  
; APPLICANT: Fillvaroff, Ellen  
; APPLICANT: Gerlicsen, Mary  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Grimaldi, Christopher  
; APPLICANT: Gurney, Austin  
; APPLICANT: Hillian, Kenneth  
; APPLICANT: Klijavlin, Ivar  
; APPLICANT: Napier, Mary  
; APPLICANT: Roy, Margaret  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Wood, William  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEAR  
; FILE REFERENCE: P2548PLIC1  
; CURRENT APPLICATION NUMBER: US/09/944,896  
; PRIOR APPLICATION NUMBER: 09/866,028  
; PRIOR FILING DATE: 2001-08-31  
; PRIOR APPLICATION NUMBER: 09/866,028  
; PRIOR FILING DATE: 2001-05-25  
; PRIOR APPLICATION NUMBER: 60/069,334  
; PRIOR FILING DATE: December 11, 1997  
; PRIOR APPLICATION NUMBER: 60/069,335  
; PRIOR FILING DATE: December 11, 1997  
; PRIOR APPLICATION NUMBER: 60/069,278  
; PRIOR FILING DATE: December 11, 1997  
; PRIOR APPLICATION NUMBER: 60/069,425  
; PRIOR FILING DATE: December 12, 1997  
; PRIOR APPLICATION NUMBER: 60/069,696  
; PRIOR FILING DATE: December 16, 1997  
; PRIOR APPLICATION NUMBER: 60/069,694  
; PRIOR FILING DATE: December 16, 1997  
; PRIOR APPLICATION NUMBER: 60/069,702  
; PRIOR FILING DATE: December 16, 1997  
; PRIOR APPLICATION NUMBER: 60/069,870  
; PRIOR FILING DATE: December 17, 1997  
; PRIOR APPLICATION NUMBER: 60/069,873  
; PRIOR FILING DATE: December 17, 1997  
; PRIOR FILING DATE: December 18, 1997  
; PRIOR APPLICATION NUMBER: 60/068,017  
; PRIOR FILING DATE: December 18, 1997  
; PRIOR APPLICATION NUMBER: 60/070,440  
; PRIOR FILING DATE: January 5, 1998

PRIOR APPLICATION NUMBER: 60/074,086  
 PRIOR FILING DATE: February 9, 1998  
 PRIOR APPLICATION NUMBER: 60/074,092  
 PRIOR FILING DATE: February 9, 1998  
 PRIOR APPLICATION NUMBER: 60/075,945  
 PRIOR FILING DATE: February 25, 1998  
 PRIOR APPLICATION NUMBER: 60/112,850  
 PRIOR FILING DATE: December 16, 1998  
 PRIOR APPLICATION NUMBER: 60/113,296  
 PRIOR FILING DATE: December 22, 1998  
 PRIOR APPLICATION NUMBER: 60/146,222  
 PRIOR FILING DATE: July 28, 1999  
 PRIOR APPLICATION NUMBER: PCT/US98/19330  
 PRIOR FILING DATE: September 16, 1998  
 PRIOR APPLICATION NUMBER: PCT/US98/25108  
 PRIOR FILING DATE: December 1, 1998  
 PRIOR APPLICATION NUMBER: 09/216,021  
 PRIOR FILING DATE: December 16, 1998  
 PRIOR APPLICATION NUMBER: 09/218,517  
 PRIOR FILING DATE: December 22, 1998  
 PRIOR APPLICATION NUMBER: 09/234,311  
 PRIOR FILING DATE: March 3, 1999  
 PRIOR APPLICATION NUMBER: PCT/US99/12252  
 PRIOR FILING DATE: June 22, 1999  
 PRIOR APPLICATION NUMBER: PCT/US99/21090  
 PRIOR FILING DATE: September 15, 1999  
 PRIOR APPLICATION NUMBER: PCT/US99/28409  
 PRIOR FILING DATE: No. US2002168715A1ember 30, 1999  
 PRIOR APPLICATION NUMBER: PCT/US99/28313  
 PRIOR FILING DATE: No. US200216715A1ember 30, 1999  
 PRIOR APPLICATION NUMBER: PCT/US99/28301  
 PRIOR FILING DATE: December 1, 1999  
 PRIOR APPLICATION NUMBER: PCT/US99/30095  
 PRIOR FILING DATE: December 16, 1999  
 PRIOR APPLICATION NUMBER: PCT/US00/03565  
 PRIOR FILING DATE: February 11, 2000  
 PRIOR APPLICATION NUMBER: PCT/US00/04414  
 PRIOR FILING DATE: February 22, 2000  
 PRIOR APPLICATION NUMBER: PCT/US00/05841  
 PRIOR FILING DATE: March 2, 2000  
 PRIOR APPLICATION NUMBER: PCT/US00/08439  
 PRIOR FILING DATE: March 30, 2000  
 PRIOR APPLICATION NUMBER: PCT/US00/14042  
 PRIOR FILING DATE: May 22, 2000  
 PRIOR APPLICATION NUMBER: PCT/US00/20710  
 PRIOR FILING DATE: July 28, 2000  
 PRIOR APPLICATION NUMBER: PCT/US00/32678  
 PRIOR FILING DATE: December 1, 2000  
 PRIOR APPLICATION NUMBER: PCT/US01/06520  
 PRIOR FILING DATE: February 28, 2001  
 NUMBER OF SEQ ID NOS: 120  
 SEQ ID NO: 2  
 LENGTH: 379  
 TYPE: PRT  
 ORGANISM: Homo Sapien  
 US-09-944-896-2

Query Match 100.0% Score 1992; DB 9; Length 379;  
 Best Local Similarity 100.0%; Pred. No. 8 6e-159;  
 Matches 3/9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MKEYVLLFLALCSAKPFESSPHIALKNNMILKDMEDTDDDDDDDDDDDDNSLFPTR 60  
 Db 1 MKEYVLLFLALCSAKPFESSPHIALKNNMILKDMEDTDDDDDDDDDDNSLFPTR 60

Qy 61 PRSHFFPDLPMPCPFGQCYSRVHCSIDLGTSYPTNIPFDTRMLDQNKKIKEIKEND 120  
 Db 61 PRSHFFPDLPMPCPFGQCYSRVHCSIDLGTSYPTNIPFDTRMLDQNKKIKEIKEND 120

Qy 121 FGGLTSVLGLANNKLUKTHKRAFLTKKLRLYLSHNOSEIPUNPLKSAELRHHEN 180  
 Db 121 FGGLTSVLGLANNKLUKTHKRAFLTKKLRLYLSHNOSEIPUNPLKSAELRHHEN 180

Qy 181 KVKKIQKDTFFKGMMALHVLENSANPLDNNINGPDKAEGVTVFHRIAEAKLTSVPKGGLPP 240  
 Db 181 KVKKIQKDTFFKGMMALHVLENSANPLDNNINGPDKAEGVTVFHRIAEAKLTSVPKGGLPP 240

Qy 241 TLLELHDYDKNTSTVELEDFKRYKEIQLRGNNKTIDENGSLANIPRVEIHLENNKL 300  
 Db 241 TLLELHDYDKNTSTVELEDFKRYKEIQLRGNNKTIDENGSLANIPRVEIHLENNKL 300

Qy 301 KKTPSGCLPELKYLQIIFLHSNSIARYGVNDCPYVKMKRSLYSALSIFNPKYWEMOP 360  
 Db 301 KKTPSGCLPELKYLQIIFLHSNSIARYGVNDCPYVKMKRSLYSALSIFNPKYWEMOP 360

Qy 361 ATFRCVLSRMSVQLGNFGM 379  
 Db 361 ATFRCVLSRMSVQLGNFGM 379

RESULT 4  
 US -09-944-944-2  
 ; Sequence 2, Application US/0944944  
 ; Patent No. US200217463A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Baker, Kevin  
 ; APPLICANT: Botstein, David  
 ; APPLICANT: Eaton, Dan  
 ; APPLICANT: Ferrara, Napoleone  
 ; APPLICANT: Filvaroff, Ellen  
 ; APPLICANT: Gerritsen, Mary  
 ; APPLICANT: Godowski, Paul  
 ; APPLICANT: Grimaldi, Christopher  
 ; APPLICANT: Gurley, Austin  
 ; APPLICANT: Hillian, Kenneth  
 ; APPLICANT: Kilaviv, Ivar  
 ; APPLICANT: Naylor, Mary  
 ; APPLICANT: Roy, Margaret  
 ; APPLICANT: Tumas, Daniel  
 ; APPLICANT: Wood, William  
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
 ; FILE REFERENCE: P2548P1C1  
 ; CURRENT APPLICATION NUMBER: US/09/944,944  
 ; CURRENT FILING DATE: 2001-09-26  
 ; PRIOR APPLICATION NUMBER: 09-866,028  
 ; PRIOR FILING DATE: 2001-05-25  
 ; PRIOR APPLICATION NUMBER: 60/087,411  
 ; PRIOR FILING DATE: December 3, 1997  
 ; PRIOR APPLICATION NUMBER: 60/069,334  
 ; PRIOR FILING DATE: December 11, 1997  
 ; PRIOR APPLICATION NUMBER: 60/069,335  
 ; PRIOR FILING DATE: December 11, 1997  
 ; PRIOR APPLICATION NUMBER: 60/069,278  
 ; PRIOR FILING DATE: December 16, 1997  
 ; PRIOR APPLICATION NUMBER: 60/069,425  
 ; PRIOR FILING DATE: December 12, 1997  
 ; PRIOR APPLICATION NUMBER: 60/069,696  
 ; PRIOR FILING DATE: December 16, 1997  
 ; PRIOR APPLICATION NUMBER: 60/069,694  
 ; PRIOR FILING DATE: December 16, 1997  
 ; PRIOR APPLICATION NUMBER: 60/069,702  
 ; PRIOR FILING DATE: December 16, 1997  
 ; PRIOR APPLICATION NUMBER: 60/069,870  
 ; PRIOR FILING DATE: December 17, 1997  
 ; PRIOR APPLICATION NUMBER: 60/068,017  
 ; PRIOR FILING DATE: December 18, 1997  
 ; PRIOR APPLICATION NUMBER: 60/070,440  
 ; PRIOR FILING DATE: January 5, 1998  
 ; PRIOR APPLICATION NUMBER: 60/074,086  
 ; PRIOR FILING DATE: February 9, 1998  
 ; PRIOR APPLICATION NUMBER: 60/074,092  
 ; PRIOR FILING DATE: February 9, 1998

PRIOR APPLICATION NUMBER: 60/075, 945  
 PRIOR FILING DATE: February 25, 1998  
 PRIOR APPLICATION NUMBER: 60/112, 850  
 PRIOR APPLICATION NUMBER: December 16, 1998  
 PRIOR FILING DATE: December 16, 1998  
 PRIOR APPLICATION NUMBER: 60/113, 296  
 PRIOR FILING DATE: December 22, 1998  
 PRIOR APPLICATION NUMBER: 60/114, 222  
 PRIOR FILING DATE: July 28, 1999  
 PRIOR APPLICATION NUMBER: PCT/US98/19330  
 PRIOR FILING DATE: September 16, 1998  
 PRIOR APPLICATION NUMBER: PCT/US98/25108  
 PRIOR FILING DATE: December 1, 1998  
 PRIOR APPLICATION NUMBER: 09/216, 021  
 PRIOR FILING DATE: December 16, 1998  
 PRIOR APPLICATION NUMBER: 09/218, 517  
 PRIOR FILING DATE: December 22, 1998  
 PRIOR APPLICATION NUMBER: 09/254, 311  
 PRIOR FILING DATE: March 3, 1999  
 PRIOR APPLICATION NUMBER: PCT/US99/12252  
 PRIOR FILING DATE: June 22, 1999  
 PRIOR APPLICATION NUMBER: PCT/US99/21090  
 PRIOR FILING DATE: September 15, 1999  
 PRIOR APPLICATION NUMBER: PCT/US99/28409  
 PRIOR FILING DATE: No. US2002173463Aember 30, 1999  
 PRIOR APPLICATION NUMBER: PCT/US99/28313  
 PRIOR FILING DATE: No. US2002173463Aember 30, 1999  
 PRIOR APPLICATION NUMBER: PCT/US99/28301  
 PRIOR FILING DATE: December 1, 1999  
 PRIOR APPLICATION NUMBER: PCT/US99/30095  
 PRIOR FILING DATE: December 16, 1999  
 PRIOR APPLICATION NUMBER: PCT/US00/03565  
 PRIOR FILING DATE: February 11, 2000  
 PRIOR APPLICATION NUMBER: PCT/US00/04414  
 PRIOR FILING DATE: February 22, 2000  
 PRIOR APPLICATION NUMBER: PCT/US00/05841  
 PRIOR FILING DATE: March 2, 2000  
 PRIOR APPLICATION NUMBER: PCT/US00/08439  
 PRIOR FILING DATE: March 30, 2000  
 PRIOR APPLICATION NUMBER: PCT/US00/14042  
 PRIOR FILING DATE: May 22, 2000  
 PRIOR FILING DATE: July 28, 2000  
 PRIOR APPLICATION NUMBER: PCT/US00/32678  
 PRIOR FILING DATE: December 1, 2000  
 PRIOR APPLICATION NUMBER: PCT/US01/06520  
 PRIOR FILING DATE: February 28, 2001  
 NUMBER OF SEQ ID NOS: 120  
 SEQ ID NO 2  
 LENGTH: 379  
 TYPE: PRT  
 ORGANISM: Homo Sapien  
 US-09-944-94-2

Query Match 100.0%; Score 1992; DB 9; Length 379;  
 Best Local Similarity 100.0%; Pred. No. 8. 6e-15; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MKEYVLLFLALCSAKPFFSPSHIALKNNMLKDMEDTDDDDDDDDDDNSLFPTR E 60  
 Db 1 MKEYVLLFLALCSAKPFFSPSHIALKNNMLKDMEDTDDDDDDDDDDNSLFPTR E 60

Qy 61 PRSHFFPFDLFPMCPFGCQCYSRVYHCSDLGLTSYPTNLPFDTRMLDQNKKKEIKEND 120  
 Db 61 PRSHFFPFDLFPMCPFGCQCYSRVYHCSDLGLTSYPTNLPFDTRMLDQNKKKEIKEND 120

Qy 121 FGLTSLYGLILNNNNKLTKIHPKAFLTTPKLRYLQLHNLSEIPLNPKSLAELRTHEN 180  
 Db 121 FGLTSLYGLILNNNNKLTKIHPKAFLTTPKLRYLQLHNLSEIPLNPKSLAELRTHEN 180

Qy 181 KVKKIQKDTEPKGMNALHVLEMSANPLDNGLEPGAFEGTVFHRIAEAKLTSVPKGLLP 240  
 Db 181 KVKKIQKDTEPKGMNALHVLEMSANPLDNGLEPGAFEGTVFHRIAEAKLTSVPKGLLP 240

Qy 301 KKIPSGLPELKYLQIIFLHSNSIARYGVNDFCPTVPKMKSLYSALSFFNNPVKYWEMQP 360  
 Db 301 KKIPSGLPELKYLQIIFLHSNSIARYGVNDFCPTVPKMKSLYSALSFFNNPVKYWEMQP 360

Qy 241 TLLFLHDYKNK1STVELEDFKRYKEOLRGLGNNK1TDIENGSLANIPRVREITHENNKL 300  
 Db 241 TLLFLHDYKNK1STVELEDFKRYKEOLRGLGNNK1TDIENGSLANIPRVREITHENNKL 300

Qy 301 KKIPSGLPELKYLQIIFLHSNSIARYGVNDFCPTVPKMKSLYSALSFFNNPVKYWEMQP 360  
 Db 301 KKIPSGLPELKYLQIIFLHSNSIARYGVNDFCPTVPKMKSLYSALSFFNNPVKYWEMQP 360

RESULT 5  
 US-09-944-907-2  
 ; Sequence 2, Application US-09944907  
 ; Publication No. US20020198147A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Baker, Kevin  
 ; APPLICANT: Bottstein, David  
 ; APPLICANT: Eaton, Dan  
 ; APPLICANT: Ferrara, Napoleone  
 ; APPLICANT: Filvaroff, Ellen  
 ; APPLICANT: Gerritsen, Mary  
 ; APPLICANT: Goddard, Audrey  
 ; APPLICANT: Godowski, Paul  
 ; APPLICANT: Grimaldi, Christopher  
 ; APPLICANT: Gurney, Austin  
 ; APPLICANT: Hillman, Kenneth  
 ; APPLICANT: Klijavinc, Ivar  
 ; APPLICANT: Napier, Mary  
 ; APPLICANT: Roy, Margaret  
 ; APPLICANT: Tunas, Daniel  
 ; APPLICANT: Wood, William  
 ; TITLE OF INVENTION: SECRET AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC ACIDS ENCODING THE SAME  
 ; FILE REFERENCE: P254 SP1C1  
 ; CURRENT APPLICATION NUMBER: US-09-944, 907  
 ; CURRENT FILING DATE: 2001-08-31  
 ; PRIOR APPLICATION NUMBER: 09-866, 028  
 ; PRIOR FILING DATE: 2001-05-25  
 ; NUMBER OF SEQ ID NOS: 120  
 ; SEQ ID NO 2  
 ; LENGTH: 379  
 ; TYPE: PRT  
 ; ORGANISM: Homo Sapien  
 US-09-944-907-2

Query Match 100.0%; Score 1992; DB 9; Length 379;  
 Best Local Similarity 100.0%; Pred. No. 8. 6e-15; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MKEYVLLFLALCSAKPFFSPSHIALKNNMLKDMEDTDDDDDDDDDDNSLFPTR E 60  
 Db 1 MKEYVLLFLALCSAKPFFSPSHIALKNNMLKDMEDTDDDDDDDDDDNSLFPTR E 60

Qy 61 PRSHFFPFDLFPMCPFGCQCYSRVYHCSDLGLTSVNPFDTRMLDQNKKKEIKEND 120  
 Db 61 PRSHFFPFDLFPMCPFGCQCYSRVYHCSDLGLTSVNPFDTRMLDQNKKKEIKEND 120

Qy 121 FGLTSLYGLILNNNNKLTKIHPKAFLTTPKLRYLQLHNLSEIPLNPKSLAELRTHEN 180  
 Db 121 FGLTSLYGLILNNNNKLTKIHPKAFLTTPKLRYLQLHNLSEIPLNPKSLAELRTHEN 180

Qy 181 KVKKIQKDTEPKGMNALHVLEMSANPLDNGLEPGAFEGTVFHRIAEAKLTSVPKGLLP 240  
 Db 181 KVKKIQKDTEPKGMNALHVLEMSANPLDNGLEPGAFEGTVFHRIAEAKLTSVPKGLLP 240

Qy 241 TLLFLHDYKNK1STVELEDFKRYKEOLRGLGNNK1TDIENGSLANIPRVREITHENNKL 300  
 Db 241 TLLFLHDYKNK1STVELEDFKRYKEOLRGLGNNK1TDIENGSLANIPRVREITHENNKL 300

Qy 301 KKIPSGLPELKYLQIIFLHSNSIARYGVNDFCPTVPKMKSLYSALSFFNNPVKYWEMQP 360  
 Db 301 KKIPSGLPELKYLQIIFLHSNSIARYGVNDFCPTVPKMKSLYSALSFFNNPVKYWEMQP 360

Db 301 KKIPSGLPEKYLQTLIFHSNSIARVGNDFCPTVPKKKSLYSATISLFFNNVKYWEMQP 360  
 Qy 361 ATFRCVLSRMSVQLGNFGM 379  
 Db 361 ATFRCVLSRMSVQLGNFGM 379

RESULT 6  
 US-09-944-457-2  
 Sequence 2 Application US/09444929  
 Publication No. US20020197612A1  
 GENERAL INFORMATION:  
 APPLICANT: Baker, Kevin  
 APPLICANT: Botstein, David  
 APPLICANT: Eaton, Dan  
 APPLICANT: Ferrara, Napoleone  
 APPLICANT: Filvaroff, Ellen  
 APPLICANT: Gerritsen, Mary  
 APPLICANT: Godowski, Paul  
 APPLICANT: Goddard, Audrey  
 APPLICANT: Grimaldi, Christopher  
 APPLICANT: Gurney, Austin  
 APPLICANT: Hillian, Kenneth  
 APPLICANT: Kijaviv, Ivor  
 APPLICANT: Napier, Mary  
 APPLICANT: Roy, Margaret  
 APPLICANT: Tumas, Daniel  
 APPLICANT: Wood, William

TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC ACIDS ENCODING THE SAME

FILE REFERENCE: P2448PIC1

CURRENT APPLICATION NUMBER: US/09/944, 929  
 CURRENT FILING DATE: 2001-08-31  
 PRIOR APPLICATION NUMBER: 09/866, 028  
 PRIOR FILING DATE: 2001-05-25  
 NUMBER OF SEQ ID NOS: 120  
 SEQ ID NO 2  
 LENGTH: 379  
 TYPE: PRT  
 ORGANISM: Homo Sapien  
 US-09-944-457-2

Query Match 100.0%; Score 1992; DB 9; Length 379;  
 Best Local Similarity 100.0%; Pred. No. 8.6e-159; Indels 0; Gaps 0;  
 Matches 379; Conservative

Oy 1 MKEYVLLFLALCSAKPFSPSHIAALKNMKLKDMEOTDDDDDDDDNSLPFTRE 60  
 Db 1 MKEYVLLFLALCSAKPFSPSHIAALKNMKLKDMEOTDDDDDDDDNSLPFTRE 60  
 61 PRSHFFPDLPMPGQCYSRVHQSDLGITSVPNTNPETRMLDQNKKIKEKEND 120  
 61 PRSHFFPDLPMPGQCYSRVHQSDLGITSVPNTNPETRMLDQNKKIKEKEND 120  
 Qy 121 FGKTSLYGLILNNKLTIKHKAFLTKLFLRRLYLSHNQSEIPINLPKSAELRHEN 180  
 Db 121 FGKTSLYGLILNNKLTIKHKAFLTKLFLRRLYLSHNQSEIPINLPKSAELRHEN 180  
 241 TLLELHDYNIKISTVLEDEKKYKELQPLGIGNNKTIDENGSLANIPRVFTHLENKL 300  
 Qy 181 KVVKTIQDKDFKGMMALHVLEMSANPLDNGTPEPGAFFGTVPHIRIAEAKLTSVPKGLPP 240  
 Db 181 KVVKTIQDKDFKGMMALHVLEMSANPLDNGTPEPGAFFGTVPHIRIAEAKLTSVPKGLPP 240  
 301 KIPSGLPEKYLQTLIFHSNSIARVGNDFCPTVPMKKSLYSATISLFFNNVKYWEMQP 360  
 Db 301 KIPSGLPEKYLQTLIFHSNSIARVGNDFCPTVPMKKSLYSATISLFFNNVKYWEMQP 360  
 Qy 361 ATFRCVLSRMSVQLGNFGM 379

Db 361 ATFRCVLSRMSVQLGNFGM 379  
 RESULT 7  
 US-10-028-072-328  
 Sequence 328, Application US/10028072  
 Publication No. US20030004311A1  
 GENERAL INFORMATION:  
 APPLICANT: Baker, Kevin P.  
 APPLICANT: Beresini, Maureen  
 APPLICANT: DeForge, Laura  
 APPLICANT: Desnoyers, Luc  
 APPLICANT: Filvaroff, Ellen  
 APPLICANT: Gao, Wei Qiang  
 APPLICANT: Gerritsen, Mary E.  
 APPLICANT: Goddard, Audrey  
 APPLICANT: Godowski, Paul J.  
 APPLICANT: Gurney, Austin L.  
 APPLICANT: Sherwood, Steven  
 APPLICANT: Smith, Victoria  
 APPLICANT: Stewart, Timothy A.  
 APPLICANT: Tumas, Daniel  
 APPLICANT: Watanabe, Colin K  
 APPLICANT: Wood, William  
 APPLICANT: Zhang  
 TITLE OF INVENTION:  
 FILE REFERENCE:  
 CURRENT APPLICATION NUMBER: US/10/028, 072  
 CURRENT FILING DATE: 2001-12-19  
 PRIOR APPLICATION NUMBER: 60/049911  
 PRIOR FILING DATE: 1997-06-18  
 PRIOR APPLICATION NUMBER: 60/056974  
 PRIOR FILING DATE: 1997-08-26  
 PRIOR APPLICATION NUMBER: 60/059113  
 PRIOR FILING DATE: 1997-09-17  
 PRIOR APPLICATION NUMBER: 60/059115  
 PRIOR FILING DATE: 1997-09-17  
 PRIOR APPLICATION NUMBER: 60/059117  
 PRIOR FILING DATE: 1997-09-17  
 PRIOR APPLICATION NUMBER: 60/059122  
 PRIOR FILING DATE: 1997-09-17  
 PRIOR APPLICATION NUMBER: 60/059184  
 PRIOR FILING DATE: 1997-09-17  
 PRIOR APPLICATION NUMBER: 60/059263  
 PRIOR FILING DATE: 1997-09-18  
 PRIOR APPLICATION NUMBER: 60/059352  
 PRIOR FILING DATE: 1997-09-19  
 PRIOR APPLICATION NUMBER: 60/059588  
 PRIOR FILING DATE: 1997-09-19  
 PRIOR APPLICATION NUMBER: 60/059836  
 PRIOR FILING DATE: 1997-09-24  
 PRIOR APPLICATION NUMBER: 60/062250  
 PRIOR FILING DATE: 1997-10-17  
 PRIOR APPLICATION NUMBER: 60/062285  
 PRIOR FILING DATE: 1997-10-17  
 PRIOR APPLICATION NUMBER: 60/062287  
 PRIOR FILING DATE: 1997-10-17  
 PRIOR APPLICATION NUMBER: 60/062284  
 PRIOR FILING DATE: 1997-10-24  
 PRIOR APPLICATION NUMBER: 60/062816  
 PRIOR FILING DATE: 1997-10-24  
 PRIOR APPLICATION NUMBER: 60/063127  
 PRIOR FILING DATE: 1997-10-24  
 PRIOR APPLICATION NUMBER: 60/063327  
 PRIOR FILING DATE: 1997-10-27  
 PRIOR APPLICATION NUMBER: 60/063329  
 PRIOR APPLICATION NUMBER: 60/063550  
 PRIOR FILING DATE: 1997-10-28





QY 301 KKPGSLPELKYLQIIFLHSNSIARYGVNDFCPTVPMKKSLYSALISLFNNPVKYWEMQP 360  
 Db 301 KKPGSLPELKYLQIIFLHSNSIARYGVNDFCPTVPMKKSLYSALISLFNNPVKYWEMQP 360

QY 361 ATFRCVLSRMSVQLGNFGM 379  
 Db 361 ATFRCVLSRMSVQLGNFGM 379

RESULT 10

US-10-140-470-328

; Sequence 328, Application US/10140470

; Publication No. US20030022331A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Beresini, Maureen

; APPLICANT: DeForge, Laura

; APPLICANT: Desnoyers, Luc

; APPLICANT: Filvaroff, Ellen

; APPLICANT: Gao, Wei-Qiang

; APPLICANT: Gerritsen, Mary E.

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Sherwood, Steven

; APPLICANT: Smith, Victoria

; APPLICANT: Stewart, Timothy A.

; APPLICANT: Tumas, Daniel

; APPLICANT: Watanabe, Colin K

; APPLICANT: Wood, William

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC ACIDS ENCODING THE SAME

; FILE REFERENCE: P3330RIC353

; CURRENT APPLICATION NUMBER: US/10/175,746

; CURRENT FILING DATE: 2002-06-19

; Prior Application removed - See Palm or File Wrapper

; NUMBER OF SEQ ID NOS: 550

; SEQ ID NO: 328

; LENGTH: 379

; TYPE: PRT

; ORGANISM: Homo Sapien

; US-10-175-746-328

Query Match 100.0% Score 1992; DB 9; Length 379;  
 best Local Similarity 100.0%; Pred. No. 8.6e-159;  
 Matches 379; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKEYVLLFLALCSAKPFPSSPHIAALKNMMLKDMEDTDDDDDDDDDDDDDDDDDDNSLFPTR 60  
 Db 1 MKEYVLLFLALCSAKPFPSSPHIAALKNMMLKDMEDTDDDDDDDDDDDDDDDDDDNSLFPTR 60

QY 1 MKEYVLLFLALCSAKPFPSSPHIAALKNMMLKDMEDTDDDDDDDDDDDDDDDDDDNSLFPTR 60  
 Db 1 MKEYVLLFLALCSAKPFPSSPHIAALKNMMLKDMEDTDDDDDDDDDDDDDDDDDDNSLFPTR 60

Query Match 100.0% Score 1992; DB 9; Length 379;  
 best Local Similarity 100.0%; Pred. No. 8.6e-159;  
 Matches 379; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKEYVLLFLALCSAKPFPSSPHIAALKNMMLKDMEDTDDDDDDDDDDDDDDDDDDNSLFPTR 60  
 Db 1 MKEYVLLFLALCSAKPFPSSPHIAALKNMMLKDMEDTDDDDDDDDDDDDDDDDDDNSLFPTR 60

QY 61 PRSHFFPDLPMPCPGCCYSRVRHCSLDGLTSVPNTIPFDTRMLDQMNKIKEIKEND 120  
 Db 61 PRSHFFPDLPMPCPGCCYSRVRHCSLDGLTSVPNTIPFDTRMLDQMNKIKEIKEND 120

QY 121 FGKLTSLYGLLNNKKLTKHKAFTTKKLFRYLISHNOLSEIPLNLPSLAELRHEN 180  
 Db 121 FGKLTSLYGLLNNKKLTKHKAFTTKKLFRYLISHNOLSEIPLNLPSLAELRHEN 180

QY 181 KVKKIQDKTEGMNALHYLEMANSPLDNNGIEPGAFEGTVFHIRIAEAKLTSVPKGGLPP 240  
 Db 181 KVKKIQDKTEGMNALHYLEMANSPLDNNGIEPGAFEGTVFHIRIAEAKLTSVPKGGLPP 240

QY 241 TLELHLHDYNIKISTVELEDFKRYKELQRIGLGNNTIDIEENGSLANTPVREIHLENKL 300  
 Db 241 TLELHLHDYNIKISTVELEDFKRYKELQRIGLGNNTIDIEENGSLANTPVREIHLENKL 300

QY 301 KKIPSGLPELKYLQIIFLHSNSIARYGVNDFCPTVPMKKSLYSALISLFNNPVKYWEMQP 360  
 Db 301 KKIPSGLPELKYLQIIFLHSNSIARYGVNDFCPTVPMKKSLYSALISLFNNPVKYWEMQP 360

QY 361 ATFRCVLSRMSVQLGNFGM 379  
 Db 361 ATFRCVLSRMSVQLGNFGM 379

RESULT 12 US-10-176-918-328  
Sequence 328, Application US/10176918  
Publication No. US20030027275A1  
GENERAL INFORMATION:  
APPLICANT: Baker, Kevin P.  
APPLICANT: Beresini, Maureen  
APPLICANT: DeForge, Laura  
APPLICANT: Desnoyers, Luc  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Gerritsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Sherwood, Steven  
APPLICANT: Smith, Victoria  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K.  
APPLICANT: Wood, William  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
FILE REFERENCE: P3330RIC382  
CURRENT APPLICATION NUMBER: US/10/176,918  
CURRENT FILING DATE: 2002-06-20  
PRIORITY APPLICATION removed - See File wrapper or Palm  
NUMBER OF SEQ ID NOS: 550  
SEQ ID NO 328  
LENGTH: 379  
TYPE: PRT  
ORGANISM: Homo Sapien  
US-10-176-918-328

Query Match 100.0%; Score 1992; DB 9; Length 379;  
Best Local Similarity 100.0%; Pred. No. 8. 6e-159;  
Matches 379; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKEYVLLFLALCSAKPFPSPSHALKNMKLKDMEIDDODDDDDNSLEPTRE 60  
Db 1 MKEYVLLFLALCSAKPFPSPSHALKNMKLKDMEIDDODDDDDNSLEPTRE 60  
QY 61 PRSHFFPDLPFPMCPEGCCYCRRVHCSDLGTSVPTNIPDTRMLDQNKKIKETKEND 120  
Db 61 PRSHFFPDLPFPMCPEGCCYCRRVHCSDLGTSVPTNIPDTRMLDQNKKIKETKEND 120  
QY 121 FGKLTSLGLILINNNKLTKHKAFLTKLRLYLSNQLSPIPLPKSLAELRHHEN 180  
Db 121 FGKLTSLGLILINNNKLTKHKAFLTKLRLYLSNQLSPIPLPKSLAELRHHEN 180  
QY 61 PRSHFFPDLPFPMCPEGCCYCRRVHCSDLGTSVPTNIPDTRMLDQNKKIKETKEND 120  
Db 61 PRSHFFPDLPFPMCPEGCCYCRRVHCSDLGTSVPTNIPDTRMLDQNKKIKETKEND 120  
QY 121 FGKLTSLGLILINNNKLTKHKAFLTKLRLYLSNQLSPIPLPKSLAELRHHEN 180  
Db 121 FGKLTSLGLILINNNKLTKHKAFLTKLRLYLSNQLSPIPLPKSLAELRHHEN 180  
QY 181 KVKKIQKDFFKGMMALHVLEMSANLDNNIGEFGAEFGTVFHRIAEAKLTSVPKGGLPP 240  
Db 181 KVKKIQKDFFKGMMALHVLEMSANLDNNIGEFGAEFGTVFHRIAEAKLTSVPKGGLPP 240  
QY 241 TLLPHLDYKLISTVLEDFKRYKELQRGLGNNTIDIEGSLANIPVREIHLENNKL 300  
Db 241 TLLPHLDYKLISTVLEDFKRYKELQRGLGNNTIDIEGSLANIPVREIHLENNKL 300  
QY 301 KKPSSGLPELKYLQIITLHSNSTARGVYNDCEPVPKMKKSLSYASISLNNPYKVWEMQP 360  
Db 301 KKPSSGLPELKYLQIITLHSNSTARGVYNDCEPVPKMKKSLSYASISLNNPYKVWEMQP 360  
QY 361 ATFRCVLSRMSVQLGNFGM 379  
Db 361 ATFRCVLSRMSVQLGNFGM 379  
QY 301 KKPSSGLPELKYLQIITLHSNSTARGVYNDCEPVPKMKKSLSYASISLNNPYKVWEMQP 360  
Db 301 KKPSSGLPELKYLQIITLHSNSTARGVYNDCEPVPKMKKSLSYASISLNNPYKVWEMQP 360

RESULT 13 US-10-176-918-328  
Sequence 328, Application US/10176921  
Publication No. US20030027276A1  
GENERAL INFORMATION:

Page 11



SEARCHER:	PHONE:	SEARCH DATE:
VENDORS/COST (where applicable)	STN:	TYPE OF SEARCH:
SPECIFICATIONS:	AA SEDUCENCES:	NA SEDUCENCES:
DRILLS:	STRUCTURES:	BIBLIOGRAPHIC:
QUESTEI/ORBIT:	LITIGATION:	LINKS:
OTHER:	DATE PICKED UP:	DATE COMPLETED:
SEARCHER REVIEW:	SEARCHER PREP/REVIEW:	SEARCHER TIME:
OTHER:	PATENT FAMILY:	OTHER:
WWW/INTERNET:	SEQUENCE SYNS.:	SEQUENCE SYNS.:
LEXIS/LEXIS:	LEXIS/NEXIS:	LEXIS/NEXIS:
OTHER:	FULL TEXT:	OTHER:
CLERICAL:	SEARCHER PREP/REVIEW:	SEARCHER TIME:
OTHER:	OTHER:	OTHER:

Ruixiang Li  
GAU 1646  
CMI 10E18  
Mail Box 10D19  
306-0282

Thank you very much!

(iii) a standard search on SEQ ID NO: 2 against Pendiing interference amino acid databases, print out the summary table only, with 500 hits and no alignments.

(i) a standard search on SEQ ID NO: 2 against commercial amino acid databases;

Please do:

Subject: Sepulchre search of Application NO: 09/944,457  
To: STIC-Biotech/Chemlib  
From: Li, Ruixiang  
Sent: Friday, May 23, 2003 1:46 PM

STC-BioTech/ChemLib

